ABSTRACT

Title of Document: A POSITIVE MODEL OF ROUTE CHOICE BEHAVIOR AND VALUE OF TIME CALCULATION USING LONGITUDINAL GPS SURVEY DATA

Cory Krause, Master of Science, 2012

Directed By: Assistant Professor Lei Zhang
Department of Civil and Environmental Engineering

This thesis approaches the topic of value of time calculation and route choice behavior with a new and innovative methodology using a survey dataset that was uniquely designed and implemented for this purpose. The survey is a 70 day, 218 participant GPS travel survey used to track individual location constantly at one minute intervals. Using a positive behavior theory framework, an in depth knowledge database for each user is created that iteratively updates the learned behavior and experienced travel conditions for each trip the user takes. A new approach for calculating value of time is presented; using the cost and trip duration of previous trips. The bounds (or caps and floors) are averaged to achieve the individual's value of time based upon their route (and therefore cost) decisions. Also using this updating knowledge base, route decision rules are derived using machine learning algorithms to tell why a user has decided to take the toll road option for certain days, and under what conditions the user will not take the toll road option. The final contribution is a model that fully takes advantage of longitudinal GPS data to create an adaptive system for value of time calculation and positive route decision making.

A POSITIVE MODEL OF ROUTE CHOICE BEHAVIOR AND VALUE OF TIME CALCULATION USING LONGITUDINAL GPS SURVEY DATA

By

Cory Krause

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2012

Advisory Committee:
Professor Lei Zhang, Chair
Professor Ali Haghani
Professor Paul Schonfeld

Dedication

This thesis is dedicated to my Mom; who showed me the value of hard work, and my Dad; who taught me to put life into perspective.

Acknowledgements

First and foremost, I would like to acknowledge my advisor, Lei Zhang, for his assistance and guidance academically as well as in my research endeavors. Without his help and encouragement, this work would not be possible.

I wish to thank Chenfeng Xiong, Yijing Lu, Xiang He, Mostafa Mollanejad, Arefeh Nasri, Jean-Michel Tremblay, Nayel Urena, Michael Maness and Nick Ferrari for helping me in countless ways with my research work over the past few years.

Additional gratitude goes to everyone in the research lab that helped format, charge, and ship GPS devices during the frantic days before the ICC opening.

Special thanks go to my high school interns and undergraduate research assistants who helped administer the survey: Shalisa Thompson, Alecia Gibbs, Othman Chebli, and Ana Argueta.

Finally, my advisory committee members Dr. Haghani and Dr. Schonfeld have been a great help throughout my graduate (and undergraduate) education. I am very grateful for their support.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

A user of the transportation network has a basic knowledge and understanding of the network around them. This knowledge is constantly evolving on a day to day and route by route basis. The interesting aspect of the individual's knowledge is that this knowledge and belief system is entirely subjective and dependent upon historic interaction with the network. By using an individual agents' interaction data with the network (GPS data), it is possible to derive sets of rules that these agents use in order to determine on what days, which route should be taken. The agent's route choice is very much an indicator of the daily value of time, as well as being true vice versa. This is why it is important to have a single theoretical framework that calculates value of time on a day-to-day and agent-by-agent basis based on their real world interaction with the network and route decisions made in the past, as well as base their route choice rules on the calculation of value of time and other travel behavior aspects such as time, cost, and other route characteristics.

The theoretical framework for how these aspects are integrated is shown below followed by a step by step explanation of each module.

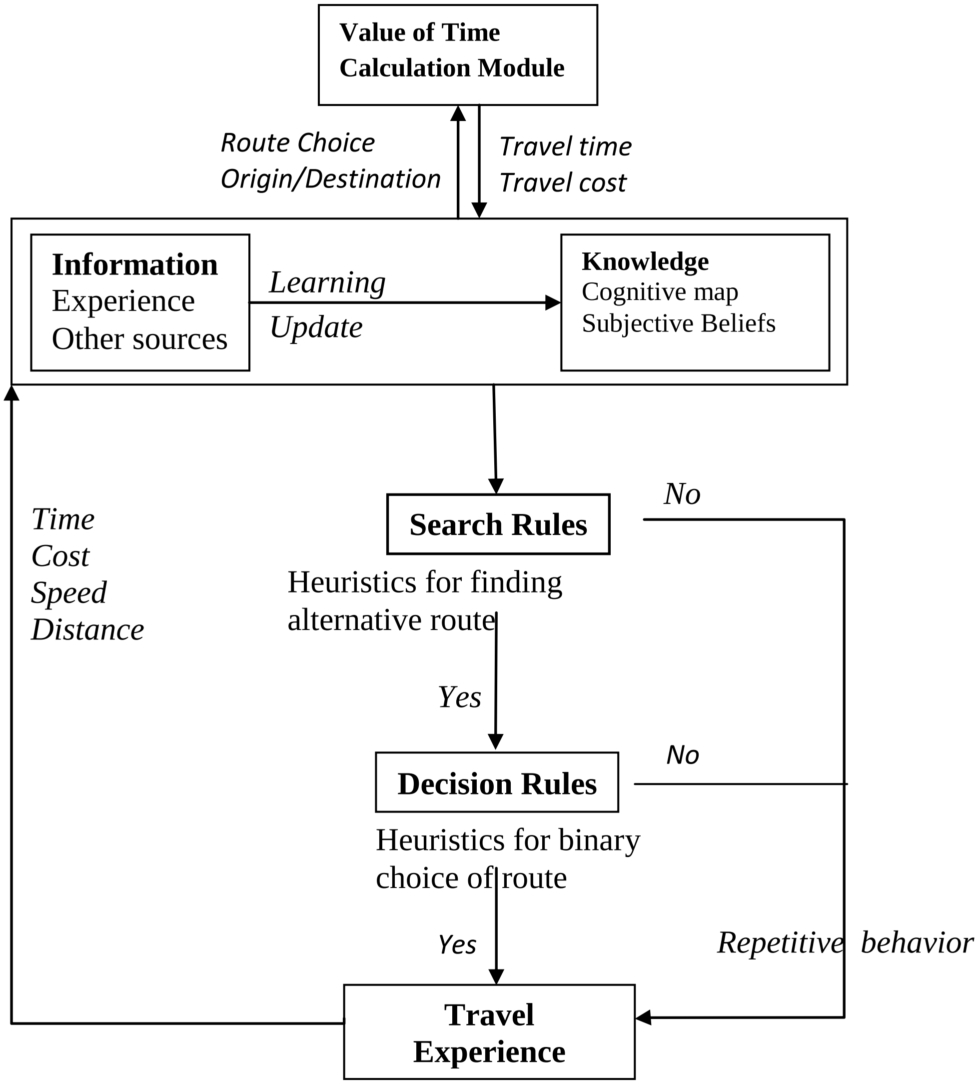


Figure 1: VOT calculation and route choice modeling framework

First, the user has pre-existing knowledge of the surrounding network and route structure. This includes basic knowledge of time and cost for the routes on their network. This information does not have to be accurate, as it may very well be biased and incorrect information derived from limited knowledge. This basic level of knowledge is an important aspect of this research in relation to the data used. It was important to acquire data at the beginning of the opening of a new facility so that the pre-existing knowledge of the individual is minimal and the learning behavior can be viewed longitudinally starting from this limited knowledge. The minimal level of knowledge that each individual begins with is that there is a new toll road in the area, and its location. All other aspects of the roadway and its' effects on travel behavior were assumed on the individual level. This includes how much time will be saved dependent upon the location of the origin and destination (this will be explained in depth in the VOT calculation module later in the thesis). The first module in the theoretical framework is the searching for alternative routes from the standard route for a set of Origin/Destination pairs. The individual user needs to determine if searching for routes is even a viable option for them. At the start of the travel survey, all users should at least be considering the alternate toll road as an option, even if they decide not to take the route. For this work, only the users that at least search for the toll road are included in the analysis. Of the 218 individuals in the dataset, only 58 of them consider the ICC toll road as a viable option. The remainder of the 160 individuals in the survey were excluded. The search rules are based upon every aspect of the user's information. While each user initially is searching for the ICC toll road option, the rule set uses the information from the individual to determine when they are satisfied with their route choice and when to stop searching (Again, this will be explained in depth in the Searching Module portion of the thesis). Once this initial searching for the alternative route is done, the decision rules come into play. The user determines via a set of rules which route to take if they have indeed decided to search in the previous step. This rule set is dependent upon true route characteristics and how the different routes match up against one another. Due to the relatively limited knowledge that the user currently has, the information available to them at the start of the decision process is minimal, and primarily consists of the perceived advantages of the new route and the real world knowledge of the previous route. In the results section of the route choice decision rules, it can be seen that much of the user's decisions are based upon perceived savings of the toll road route. Over a longer survey trial, or one whose rules are broken up dependent upon increasing survey time, it may show that the user's decision factors are different from the beginning of the trial when they had limited knowledge.

After this decision is made on whether or not to search for alternative routes and the user decides on which trip route to take, the user finally experiences their travel. The key components to their travel experience are cost and time for that particular route, but a wide range of other factors are also recorded including: average speed, distance, percent of trip (distance) that occurred on the toll road, number of segments traveled on toll road, and time of day traveled. Based on these data calculated for each trip, another set of variables is calculated by person, by OD, and by route. These statistics include: average time, change in distance from previous days, change in travel time from previous days, change in value of time calculated from previous days, change in cost, change in speed, etc... This set of variables is basically used to see how the most recent route taken compares to others taken by the same person from the same OD. Once this experience is internalized, the individual's knowledge of the network and route is more robust. This process is then repeated, with each step the knowledge of the information is slightly more complete (subjectively) and hence their searching and decision behavior is constantly improving.

Once knowledge is updated from that travel experience, after the search and route choice modules, the calculation of value of time is calculated from this knowledge base. The way that this is done, is that the knowledge of each user is stored and compared to the decision made by that user for the previous day. The start of each user's knowledge is the simple time and cost for their normal route for the origin/destination sets. After each iteration, that trip's behavior experience is recorded in the route category, such that during the information and knowledge step, the user has a database including characteristics of each origin/destination pair route. The knowledge database used for the VOT calculation for each user after roughly 10 days of travel may look something like:

Table 1: Example user knowledge base

	OD1 route 1	OD1 route 2	OD2 route 1	OD 2 route 2
Day 1	15 (min), $0			
Day 2	17 min, $0			
Day 3		12 min, $2		
Day 4	14 min, $0			
Day 5			42 min,$0	
Day 6			44 min,$0	
Day 7		13 min, $2		
Day 8				32 min, $1.2
Day 9	18 min,$0			
Day 10			39 min,$0	

After this value of time is calculated, the result is put back in the user's knowledge database to be used for deriving subsequent rules for searching and route decision making. Finally, one cycle is complete, and the search module can begin with fully updated knowledge.

1.1 Summary of Contributions

The results achieved through this research advance the field in a number of dimensions. First, the calculation of value of time gains a brand new approach using longitudinal GPS survey data. Most importantly, this estimation for value of time is the first to be able to adapt to daily changes in intra- and inter-personal value of time. Never before has the estimation for value of time been calculable on the daily and trip by trip basis. Additionally, the positive route choice behavior rules are the first of its kind to derive heuristics starting from a true limited knowledge base case and expand to a larger knowledge base. Since the GPS survey acquires data prior to the toll road opening and continues through the first 10 weeks of its use, the GPS data, and hence the heuristics describing their behavior, are able to describe truly learned behavior as travel is experienced and knowledge is gained on the toll road. The combination of these advances result in a model that accurately estimates route choice based upon constantly updated interpersonal value of time and other learned travel characteristics.

1.2 Thesis Structure

The remainder of the thesis is structured in this manner: Previous work will be explained via an in depth literature review of previous methodologies on learning route choice models and approaches for calculating value of time as well as other GPS travel surveys. The toll road that is so essential to this research will be described including the aspects that necessitated its use. The GPS survey that was implemented will be explained in depth from the inception of the pilot survey, to participant acquisition and selection, GPS specification, calibration, data management and finally coding measures used for analysis. This includes a walk-through of the website created to gain information on all participants as well as implement an independent traditional travel diary throughout the GPS survey. The calculation of value of time will be thoroughly examined and results on the average value of time including variations of value of time will be given. Changes of value of time will be shown within person and across person. The search and decision rules derived from machine learning algorithms and their impact on why users decide to change route are explained. Finally, the thesis is concluded with a summary of research done and what future work can be done using these approaches.

Chapter 2: Literature Review

This literature review covers three general fields related to this work: the positive behavior route choice modeling approach used in the search and decision modules, the value of time calculation needed for the VOT module, and finally a review of other GPS survey data and how they have been used in positive behavior modeling and VOT calculation.

2.1 Learning Route Choice Modeling:

The route choice learning model began with Horowitz (1984) using the assumption that perceived travel time was a function of the travel times in the past. The work was based on a two link system, much like the one explored in this work, where the two routes are the toll and non-toll route option. In general, three scenarios were developed from this work: 1) the perceived travel time is an average of real travel time on both routes, 2) perceived travel time (real time plus an error) average on both routes and 3) perceived travel time on the chosen route only. The problem with this approach however, is that in circumstances 1 and 2, a pre-existing knowledge of the network and actual travel times (whether perceived or true), must be known. Scenario number 3 is difficult to use however, seeing as it is difficult to estimate a perceived travel time for individuals.

Simulation models have used learning models similar to those derived in Horowitz (1984) to simulate individuals. Ben-Akiva (1991) and Emmerink (1995) have lead this research in route learning models under the assumption that users update their knowledge based upon travel experience. Both use a form of utility maximization for the selection of route based on previous knowledge.

Nakayema (2001) simulated this learning process as well, but with a different set of rules. The route choice is either based upon experience based on a limited number of past days, or experience based on all past trips. This is the first research done in the field that showed that the user's route choice is not rational and an imperfect choice can be made based upon past experiences. Arentze and Timmermans (2005) make further contributions by showing that individuals make observations about their current trip that will increase their knowledge of the environment and subsequently have an impact on later trips.

Finally, the upcoming work is based off of a system derived by Zhang (2006) in which search rules are represented via a decision tree generated from survey data (in this work, based on real world GPS data), which determines what alternative is considered. If another alternative is considered, another decision tree is implemented in order to decide whether the user will use this alternative.

2.2 Value of Time Estimation Review:

Traditionally, value of time determinations are computed using some form of multinomial logit model. Using the utility model V, the calculated value of time is shown as:

$$(\partial V / \partial TT)/(\partial V / \partial TC) \qquad (1)$$

Where TT is travel time, and TC is travel cost. This value will be reduced to the cost and time coefficients (β_{TT}/β_{TC}), which gives the marginal utility of increase by one unit in travel-time and cost. Recently work has been done to further the calculation of value of time through more complex models such as nested logit and generalized extreme value models. This work chooses to move away from these models to form a day-to-day calculation of value of time.

Traditional value of time calculations are performed not with GPS data, but with some combination of RP and SP datasets. Small et al. (2005) has looked at similar issues of calculating value of time dependent upon toll road use but with revealed and stated preference data as opposed to the GPS data used in this study. Like his previous work, the calculation of value of time is derived through mixed logit modeling through the division of the time coefficient by the cost coefficient. Cirillo & Axhausen (2006) looked into the possibility of a negative value of travel time savings using the German Mobidrive study (including 160 households and 360 individuals) in a 6 week period. The Mobidrive study included an extensive daily travel diary that focused on every aspect of an individual and household’s travel. Hess, Bierlaire, and Polak (2005) have extensively research current issues in discrete choice modeling of travel-time savings. Including in the issues is the non-zero probability of positive travel-time coefficients.

The advantage of using longitudinal GPS data is that it is possible to easily isolate the individual user’s value of time, as well as see the progression of the estimation as more travel behavior is taken into account, trip routes stabilize, and route information is learned by the user. Also, since all data is real world, it would not be possible to generate a negative value of time using the longitudinal value of time cap method. Value of time calculations have existed for several decades, starting with traditional logit models. GPS based surveys have begun to replace more traditional phone or paper travel surveys, with their advantages clearly noted. However, GPS data has not been used for the calculation of individual user's value of time.

2.3 GPS Survey Review:

Similar GPS studies to the Maryland study have looked into the idea of commute route choice. In phase I the Commute Atlanta project collected data on 182 drivers for a period of ten days and attempted to describe real world travel via a binary logit model of commuter's route choice. Findings showed an increase in multiple routes from increased arrival time flexibility. Also, an increase in idle stops may increase number of routes, with average speed increases reducing trip routes. (Li et al, 2005). In Phase II, 130 participants had car chips installed in their car for one year to track time, speed, distance, but is unable to track vehicle location. Using this dataset an in depth research was done on VMT changes due to a distance-based congestion pricing scheme. Changes to demographics such as income, household structure, work status, home location, and vehicle ownership changes were recorded and their impact to VMT was measured. While advancing the state of the art practice in GPS longitudinal study, the Commute Atlanta datasets have yet to be used in a value of time calculation.

The Microsoft Company has also looked into using real world GPS data with a large scale GPS study. Their 700 GPS device dataset has been used to build a system in which GPS points are aggregated to build a dependable road network, predict vehicle location, investigate location privacy, and determine mode choice. They did not collect travel information such as toll cost however, as would be necessary for value of time calculations.

Chapter 3: Data

In this section, the entire scale of the GPS survey conducted for this research will be explained. All data used in this research was acquired through the GPS survey spanning from October 2011 to February 2012. The breakdown of the section will be:

-The ICC toll road

-Pilot Survey

-GPS device selection

-Postcard and Email Campaign

-Website Design

-Survey Design

-Respondent Sampling

-GPS mailing and returns

-Online Travel Diaries

-Data analysis

-Survey Statistics

For ease of reading all forms sent out to participants with instructions on GPS installation and online trip diary example are in the appendix section B.

3.1 The ICC Toll Road

MD-200, also known as the InterCounty Connector (ICC), is the first all-electronic toll road to be implemented in the state of Maryland. All tolls are collected at highway speeds using the EZPass and Video Toll Rate systems. This means there is no delay in accessing the roadway for users. With the EZPass device in your car, you simply drive on the roadway as you would any other road in the area. The ICC opened on November 22nd connecting I-270 to I-95, with an additional segment planned to reach US-1 in 2014.

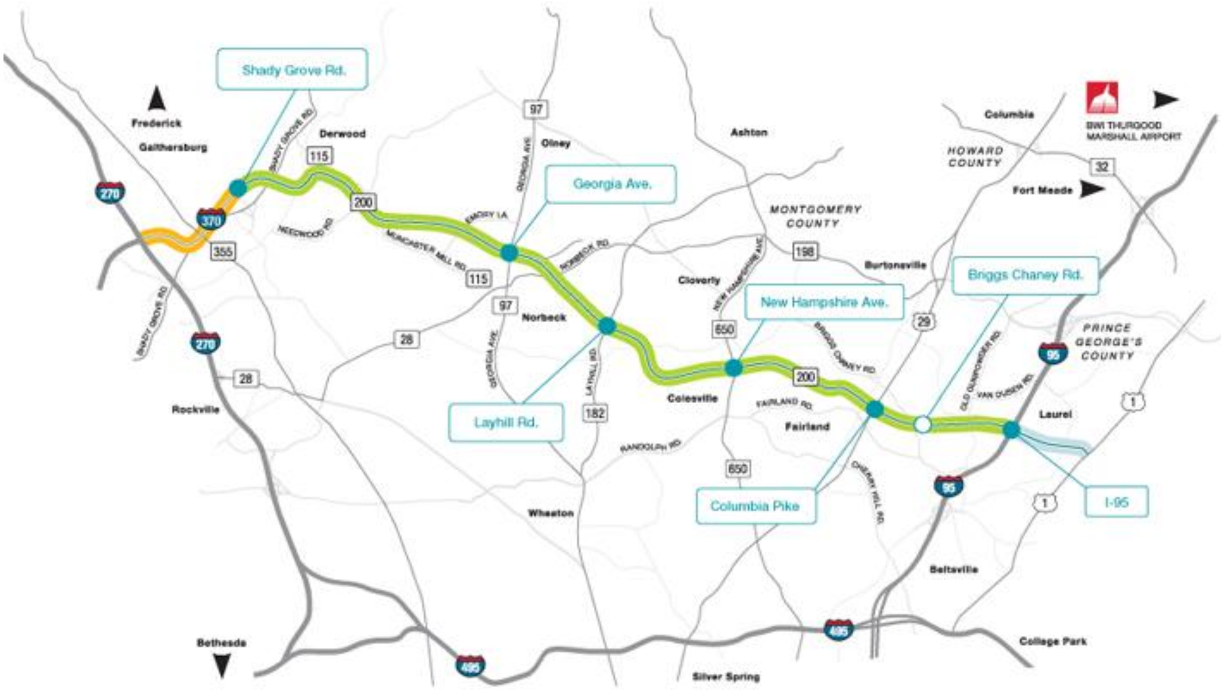


Figure 2: ICC toll road map

For practical application, the roadway was broken down into 6 sections for use in this research. The first section (in yellow) connecting I-270 and Shady Grove is a free to drive facility, and was not considered in this research. The light blue portion (east) is to be added in 2014 and was, of course, not included. Due to the nature of the research as related to learning and knowledge base of the user, it was extremely important to acquire the user's base knowledge of the network before the roadway was to be opened and begin tolling. The start date of the GPS survey was implemented on November 16th, giving ample opportunity for the individual participants to drive on the network before the new road was opened. It should also be noted that at no time were participants told to alter their travel behavior, nor would they be rewarded for travel on the ICC; the GPS survey simply wanted to capture the behavior of road users in the area.

3.2 Pilot Survey

The pilot survey was used to determine the feasibility of a large scale survey in every aspect; response rate, data collection, man hours, total time, etc. The framework for the pilot survey follows the full scale survey except for some small changes that were made after the data received from the pilot. The pilot survey lasted 2 weeks with 20 participants.

While examples of website pages are given, to get a full view of the website, please visit http://travel-survey.org/ for the main site, and http://travel-survey.org/gps for the travel diary.

3.3 GPS Device

The GPS device was decided upon before the start of the project. Previously the research team had worked with the QSTARZ 1000XT which worked well. It has a running battery life of 42 hours after two hours of charging with more than acceptable accuracy. The cost of each device is approximately $80 including shipping costs. With the device continually plugged into participants' cars the device will stay charged permanently. Even if the car is not used for 42 straight hours, the device will switch into sleep mode which drastically conserves battery life. The GPS will record at an interval of one minute. Using this interval we achieve a number of things. The data collected is able to be used successfully in determining origins and destinations. The maximum for this interval for determining destinations is two minutes, since stops can be conducted in less than this amount of time. The lower the time interval, the more dependable the data will be. Since the GPS survey lasts for a period of two months, the one minute time interval was selected. At this rate, if the car is driven continuously for two straight months, the GPS device will be able to record every data point. While the interval could be set to a lower level, and probably collect all data for the individual, it was decided to err on the side of caution in case of a near constant use situation for some participants.

Each GPS device is set to this time interval specification and given a unique identifier for each unique individual. After each device has been manually formatted and charged they are ready to be shipped to participants acquired through the postcard and email campaign.

3.4 Postcard and E-mail Campaign

The postcard and E-mail campaign is used in order to get participants for the travel survey. The main purpose of this campaign is to get people to go to the travel survey website and fill out the initial participation form for full participation.

Figure 3: Front side of postcard

Figure 4: Back side of postcard

The target area was the communities encased in the I270-MD200-I495-I95 area. This is in both Montgomery and Prince George's County. This area was considered to be the location where most of the users of the new toll road would live, and would allow us the largest viable participation. The postcard campaign was sent to 22,000 households. Of the 22,000 postcards sent, there were 893 registrations on the website for a response rate of about 4%

While the email campaign was originally used in the pilot survey, the response rate was so low that it was not used for the full scale study. The response rates for the pilot survey email campaign are as follows:

Table 2: Email campaign tracking report numbers

Metric	Total	Rate		
Sent Messages	1000	100%		
Received Messages	895	89.50%		
Total Bounces	69	6.90%		
Soft Bounces	5	0.50%		
Hard Bounces	64	6.40%		
Undelivered	36	3.60%		
Metric	**Total**	**Total Rate**	**Unique**	**Unique Rate**

Opens	23	2.57%	22	2.46%
ClickThroughs	13	1.45%	12	1.34%
Unsubscribes	1	0.11%	1	0.11%

Of the 1000 emails sent out, 22 individuals opened the email to read it and only 8 registered to take part in the online survey. The registration rate for the email campaign comes to 0.8% compared to the 4% level of the postcard survey. Even when considering the lower cost of the email campaign per person, it was still not a feasible option.

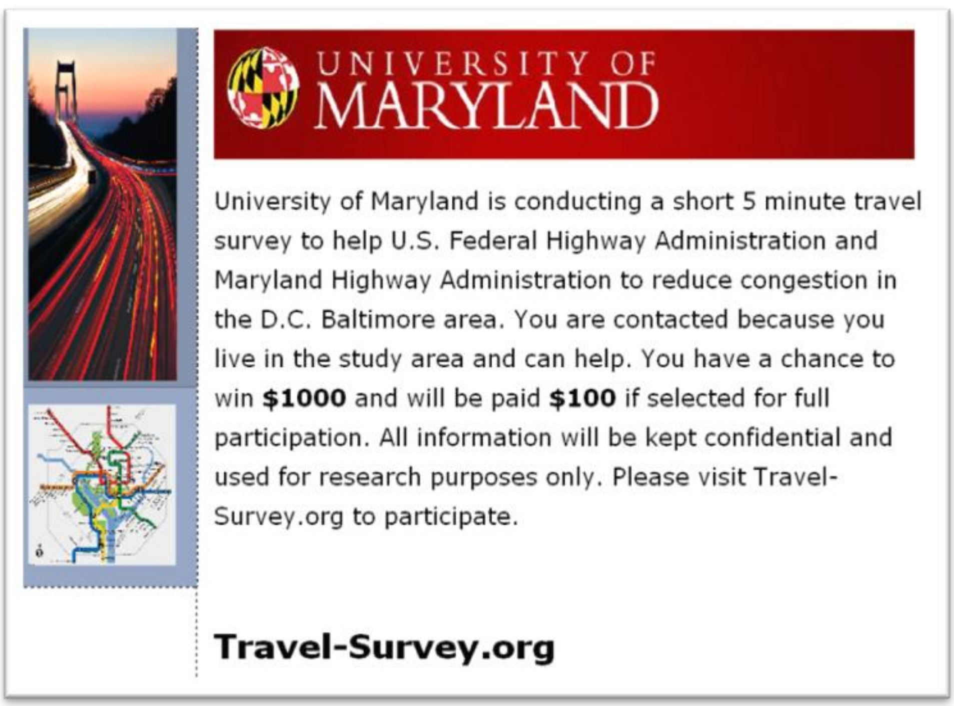


Figure 5: Email sent to individuals in the project area

- Sent out to residents of Prince Georges County, Montgomery County, and The District of Columbia.
- Mainly focuses on getting residents to go onto the website.
- Financial incentives included in order to increase response rate.

3.5 Website Design

The website is an extremely important aspect of the travel survey because we focus all possible participants to this location. In order to get people to go from receiving a postcard to participating in our GPS survey, the website must be easy to operate and understand. In this section a breakdown of each page will be shown and how it is designed to get users to participate. If viewing this document via electronic means, please visit travel-survey.org to view the website.

On each page there is a button that can be selected to take them to the participation page. From there they can fill out the initial form to allow us to contact them. The main page of the website gives a small blurb about why travel surveys are important as well as remind them of the monetary benefit of taking part in the survey. Also included on every page is a link back to the homepage, a link to the Maryland State Highway Administration (SHA), the Federal Highway Administration (FHWA), and a link to the University of Maryland (UMD) Website.

An FAQ page and contact information page are also included in case people are confused as to how to participate; we can help them along with the process. These buttons are available on every page except when filling out the participation form.

Finally, if they want to learn more about the Inter County Connector (ICC), they can access the ICC main page by clicking on the "About the ICC" link shown on every page, except for pages including forms.

Figure 6: Front page of the survey website

3.6 Survey Design

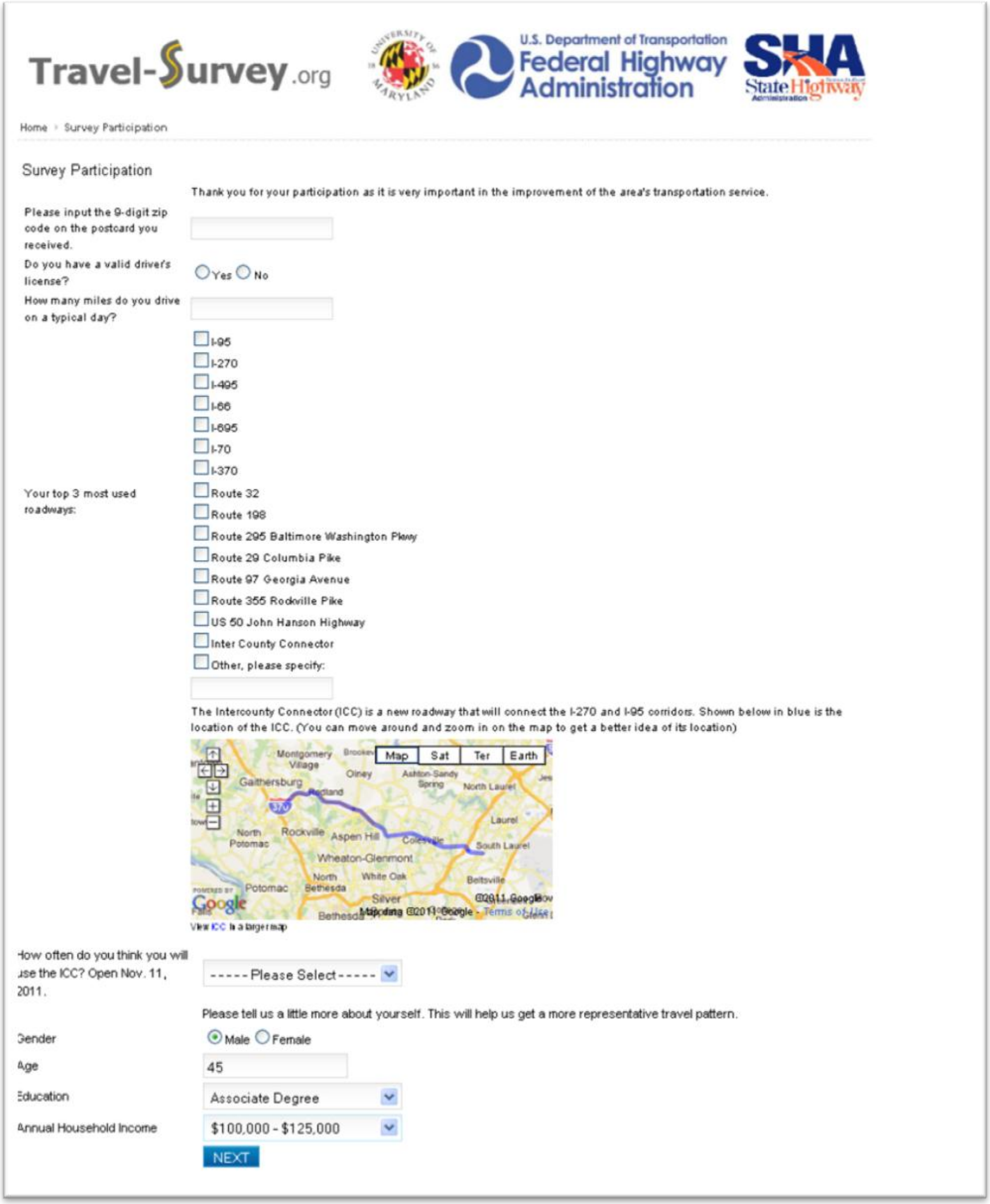
In this section the initial participation form will be explained. This form is used to get basic demographic and driving behavior information so that we can contact the participants after the respondent sampling portion. The demographic information is used for participant selection, and in later steps of the research is also used in the learning algorithms for the creation of search and decision rules.

This form is actually three separate forms: one for people who received the postcard, one for people who received an email from us, and one for those that found our website by another means. The purpose of having these different forms is for gathering slightly different information from these three groups. For example, those who got our postcard, we already have their mailing address, so it is not necessary to ask that information again. We try to ask as few questions as possible. In this way we maximize the amount of responses we get by making the survey less invasive. Below is the opening participation form screen:

Figure 7: Participation page to start the survey

Based upon which of the three buttons they click, it will take them to that form. Once the questions have started, all other links to tech support, financial supports, and participation are dropped so that people will not leave this page. The only thing left on the page is the questionnaire and the links to UMD, SHA, and FHWA.

The first page includes a series of demographic questions so that respondents can be selected for full participation and used for later analysis. Also, we gain a little knowledge about their travel patterns by asking about frequent roads taken and how likely they are to take the ICC.



Travel-Survey.org

Home › Survey Participation

Survey Participation

Thank you for your participation as it is very important in the improvement of the area's transportation service.

Please input the 9-digit zip code on the postcard you received.

Do you have a valid driver's license? Yes No

How many miles do you drive on a typical day?

Your top 3 most used roadways:

- I-95
- I-270
- I-495
- I-66
- I-695
- I-70
- I-370
- Route 32
- Route 198
- Route 295 Baltimore Washington Pkwy
- Route 29 Columbia Pike
- Route 97 Georgia Avenue
- Route 355 Rockville Pike
- US 50 John Hanson Highway
- Inter County Connector
- Other, please specify:

The Intercounty Connector (ICC) is a new roadway that will connect the I-270 and I-95 corridors. Shown below in blue is the location of the ICC. (You can move around and zoom in on the map to get a better idea of its location)

View ICC in a larger map

How often do you think you will use the ICC? Open Nov. 11, 2011. -----Please Select-----

Please tell us a little more about yourself. This will help us get a more representative travel pattern.

Gender: Male Female

Age: 45

Education: Associate Degree

Annual Household Income: $100,000 - $125,000

NEXT

Figure 8: Participation form first page

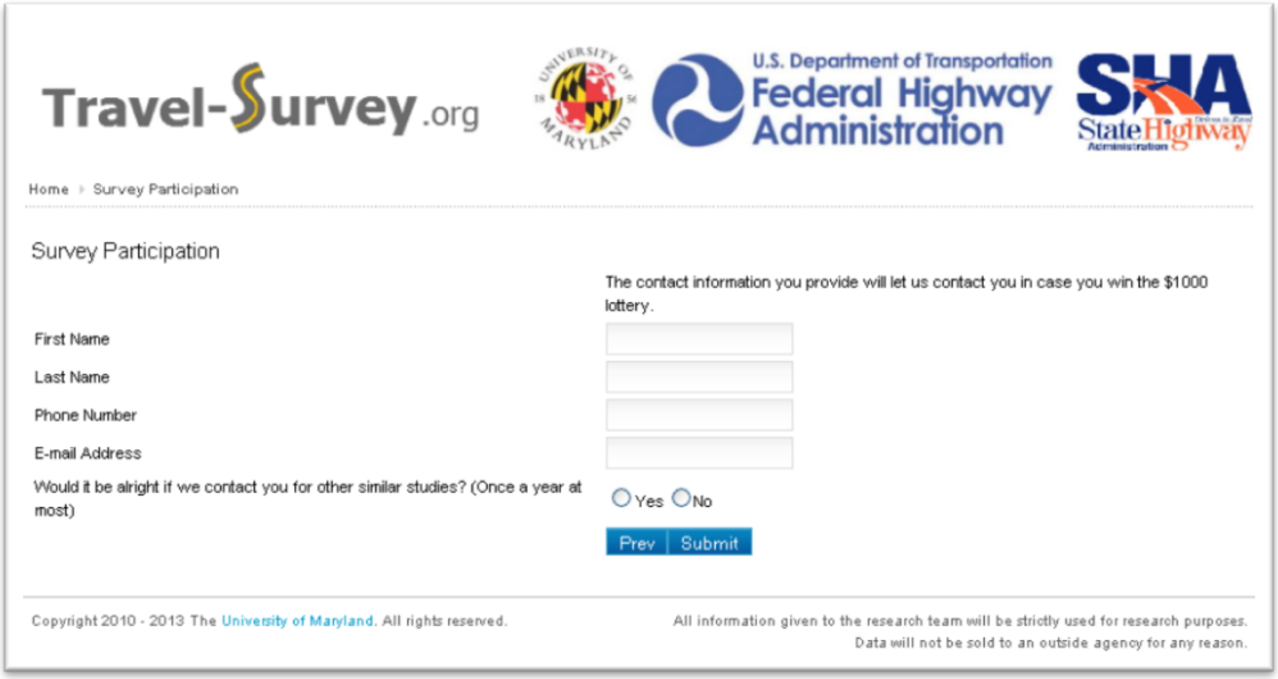


Figure 9: Participation form second page

The final page includes contact information. While users may choose to not fill out certain sections of this page, an error message comes up to make sure they meant to leave these sections blank. The section for email is mandatory however, as it serves two purposes: it allows us to contact them for the GPS portion of the survey and it is used for an instant follow-up email notifying that their submission was successful.

3.7 Respondent Sampling

Once information is gained on travel behavior and demographics, the participants for the GPS portion can then be chosen. Participants are selected in order to gain a representative sample of the surrounding area. Demographics considered are: Sex, Age, Household Income, and Education level (Appendix). While not a demographic measure, we also check that each person has a driver's license and amount of average driving per day. This is to ensure that each participant will be supplying the project with data and not simply sitting at one location not moving.

#	☐	Date Submitted	Submitter's IP Address	Language Code	Identifier	License	travel time	Most used roadways	Other Roadways	ICC	Gender	Age	Education	Income	First Name	Last Name	Phone
1	☐	2011-09-11 14:02:08	173.66.196.105	en-GB	20903-1314	Yes	25	I-495, Route 29 Columbia Pike, Other, please specify:	New Hampshire Avenue	Less than once a month	Female	53	Master's Degree	$50,000 - $75,000			
2	☐	2011-09-10 14:15:52	208.58.5.251	en-GB	20910	Yes	30	I-495, Route 29 Columbia Pike, Route 97 Georgia Avenue		Less than once a month	Female	60	Doctorate Degree	$150,000 & up			
3	☐	2011-09-05 17:25:46	70.110.22.103	en-GB	20901-1847	Yes	25	I-495, Route 29 Columbia Pike, Route 97 Georgia Avenue		Once a week	Male	47	Master's Degree	$100,000 - $125,000			
4	☐	2011-09-01 16:30:25	71.191.132.217	en-GB	20906-1118	Yes	20	I-95, I-495, Route 198		Several days a week	Male	45	Master's Degree	$150,000 & up			
5	☐	2011-08-30 16:35:48	68.50.114.31	en-GB	20906-2672	Yes	20	Route 355 Rockville Pike, Inter County Connector		Several days a week	Male	48	Associate Degree	$50,000 - $75,000			
6	☐	2011-08-30 10:49:14	216.81.81.82	en-GB	209102246	Yes	40	I-495, Route 295 Baltimore Washington Pkwy, Route 97 Georgia Avenue		Less than once a month	Male	46	Bachelor's Degree	$150,000 & up			

Figure 10: Example of participation form results

A stratified random sampling procedure was used in order to create a sample that matches demographics for the area closely. While participants were selected based upon demographic information to give the most representative sample, a bias towards over-education and high income does occur. To view the participation demographics compared to those of the surrounding area, please go to Appendix C:

3.8 GPS Mailing

Each package included the GPS device, instructions on use and the GPS charger. The devices were sent out roughly 5 days before the opening of the ICC to make sure that data was collected first of their case driving behavior. After the survey period was over (2 months) each participant received a preaddressed prepaid shipping label to be sent back to the University of Maryland. The participant simply puts the GPS device back in the box and drops it off at any FedEx location.

3.9 Travel Survey

The travel survey is an extremely important part of the GPS survey. It is the only link between the raw data from the GPS device and the actual user travel patterns. The online survey was designed to include all the pertinent travel information while being simple to fill out. On the first page the user inputs their name and the date for which they are filling out the form. This is used to link to the GPS received from the participant. Then, the next series of pages are all nearly identical, asking for information on a single trip for the day. If the user has more trips for that day, they are prompted to click the continue button, or if they are finished recording all their trips, they can finish by clicking submit. It is simple, yet provides us with time, trip purpose, travel mode, and exact location.

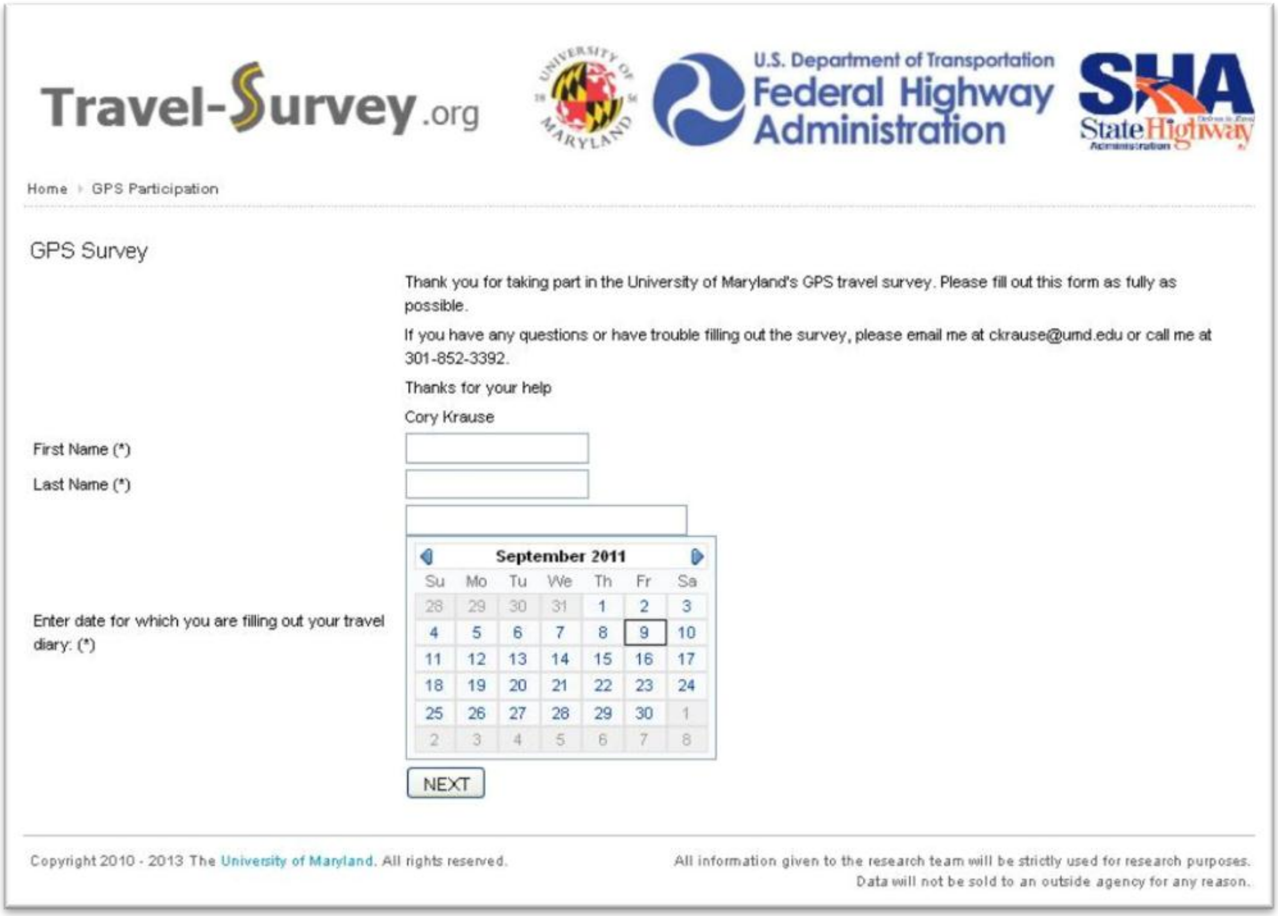


Figure 11: Travel Survey First Page

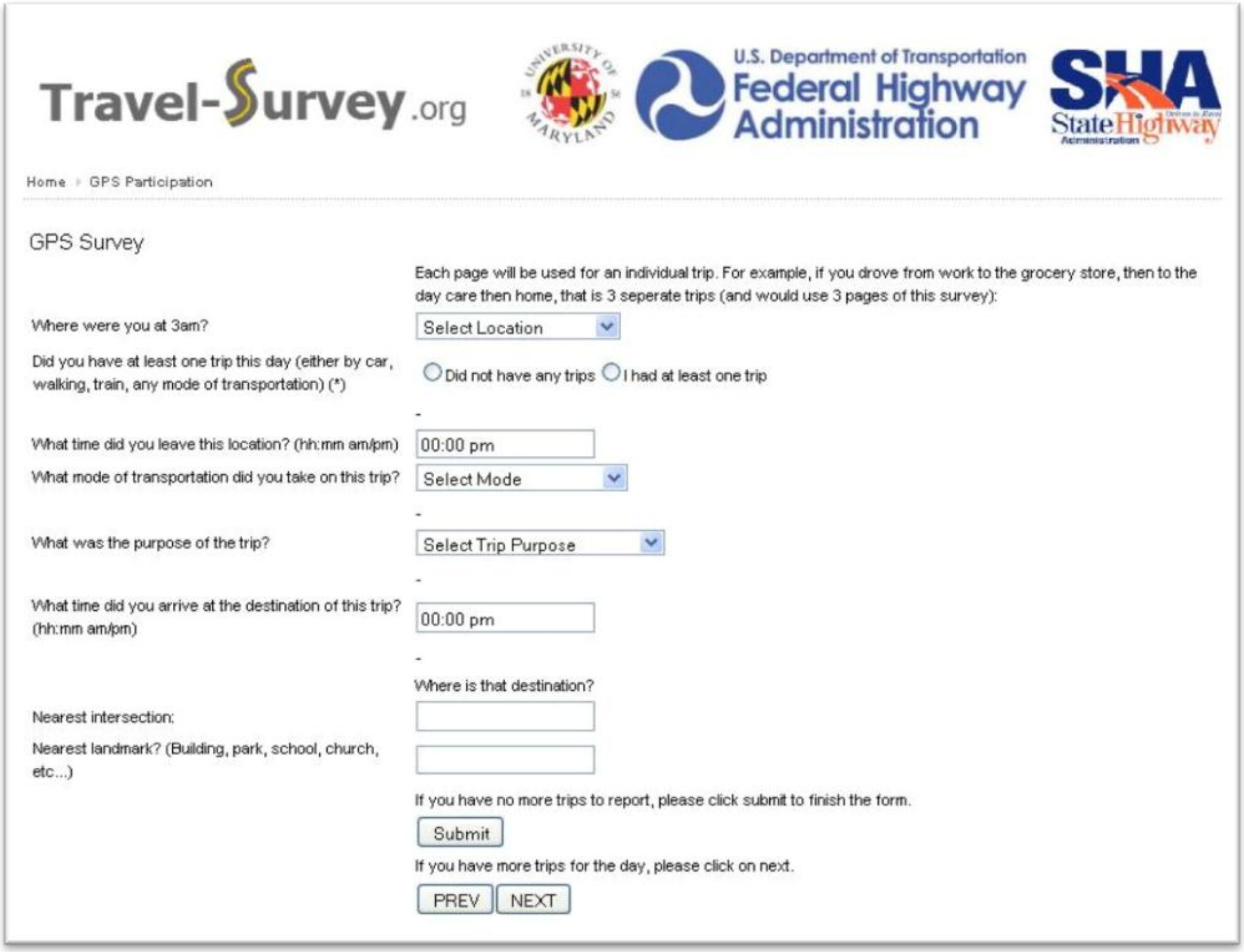


Figure 12: Travel Survey Second Page

While this information can be used to validate the GPS data, the daily trip diary was not used in this Value of Time calculation research.

3.10 GPS Data Checking

The final step in the pilot study is to upload and check the data to make sure that each device has been functioning properly. In a two-step process, this can be down fairly simply. First, open up the data in an excel file to make sure the data points are available. Then, using a geospatial tool (using GeoStats' TravTime software), make sure the data points don't show huge error in data location.

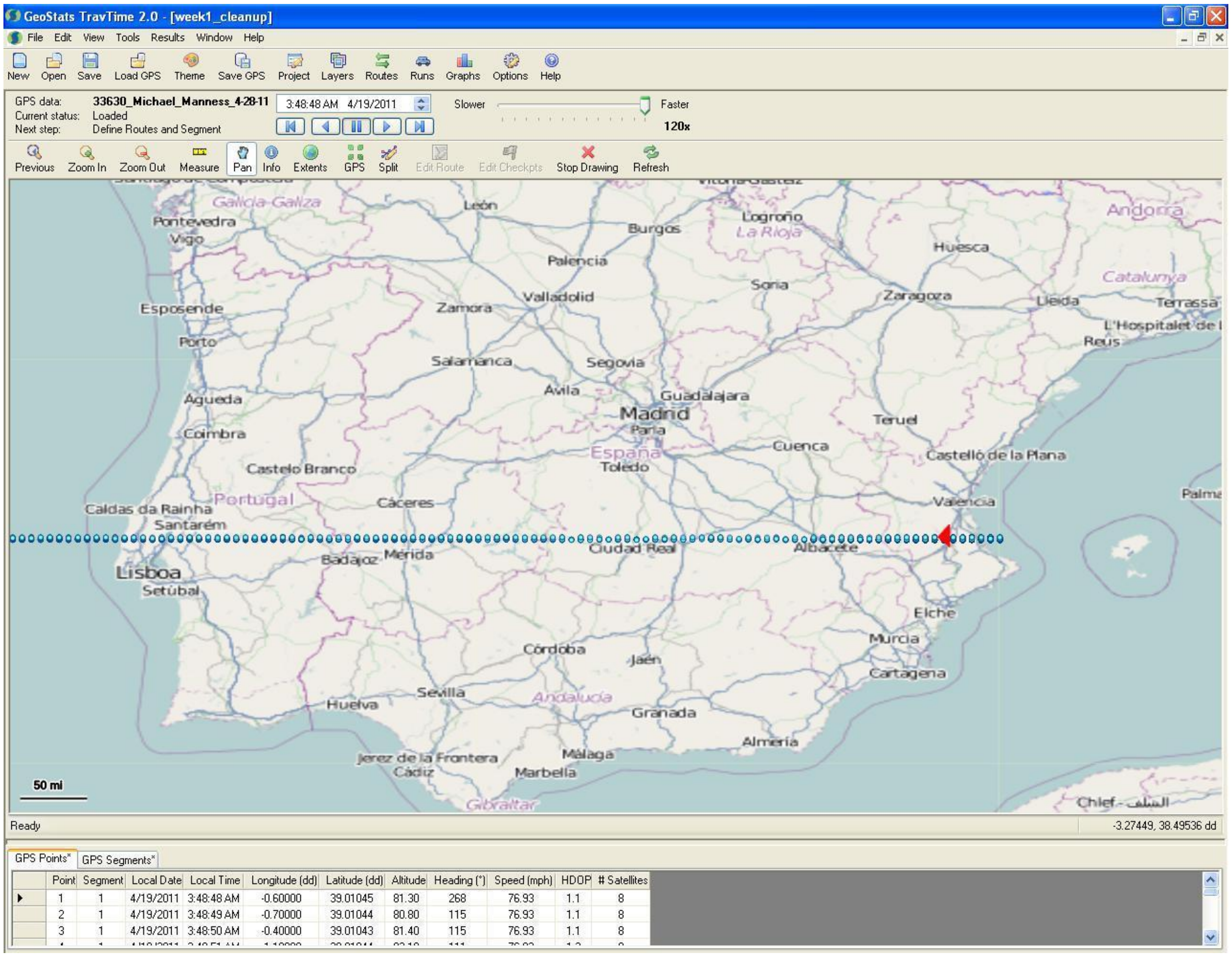


Figure 13: An example of data error.

Clearly this data is not valid seeing as it occurs on a different continent and some of the trip occurs over water. While this has occurred, the failure rate for these devices is about 5% and is bearable for the full scale survey.

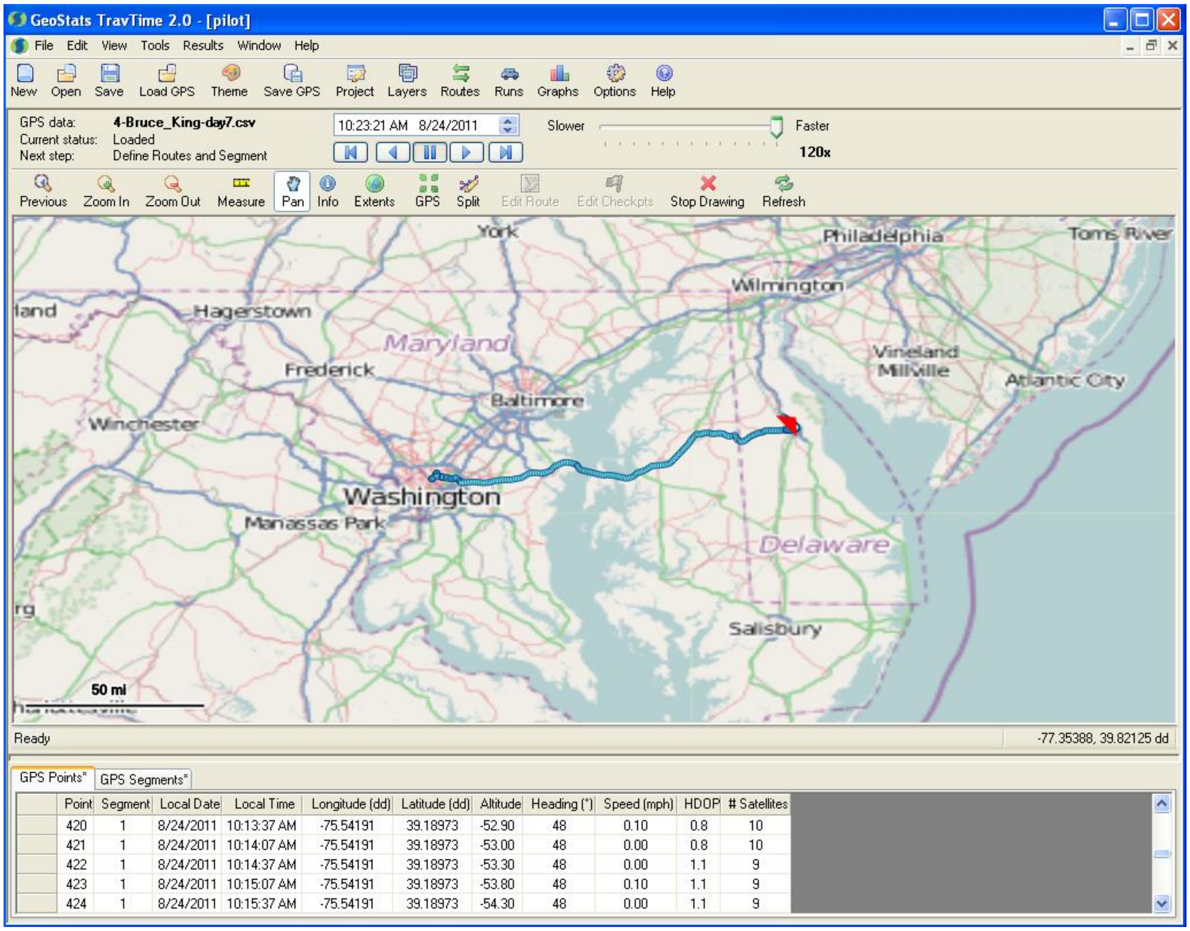


Figure 14: Correct data as recorded by the GPS device.

Above is an example of data when it correctly represents user location. Shown is a relatively long term trip from Maryland to Delaware. When zooming in (below), you can see the exact route taken by the traveler.

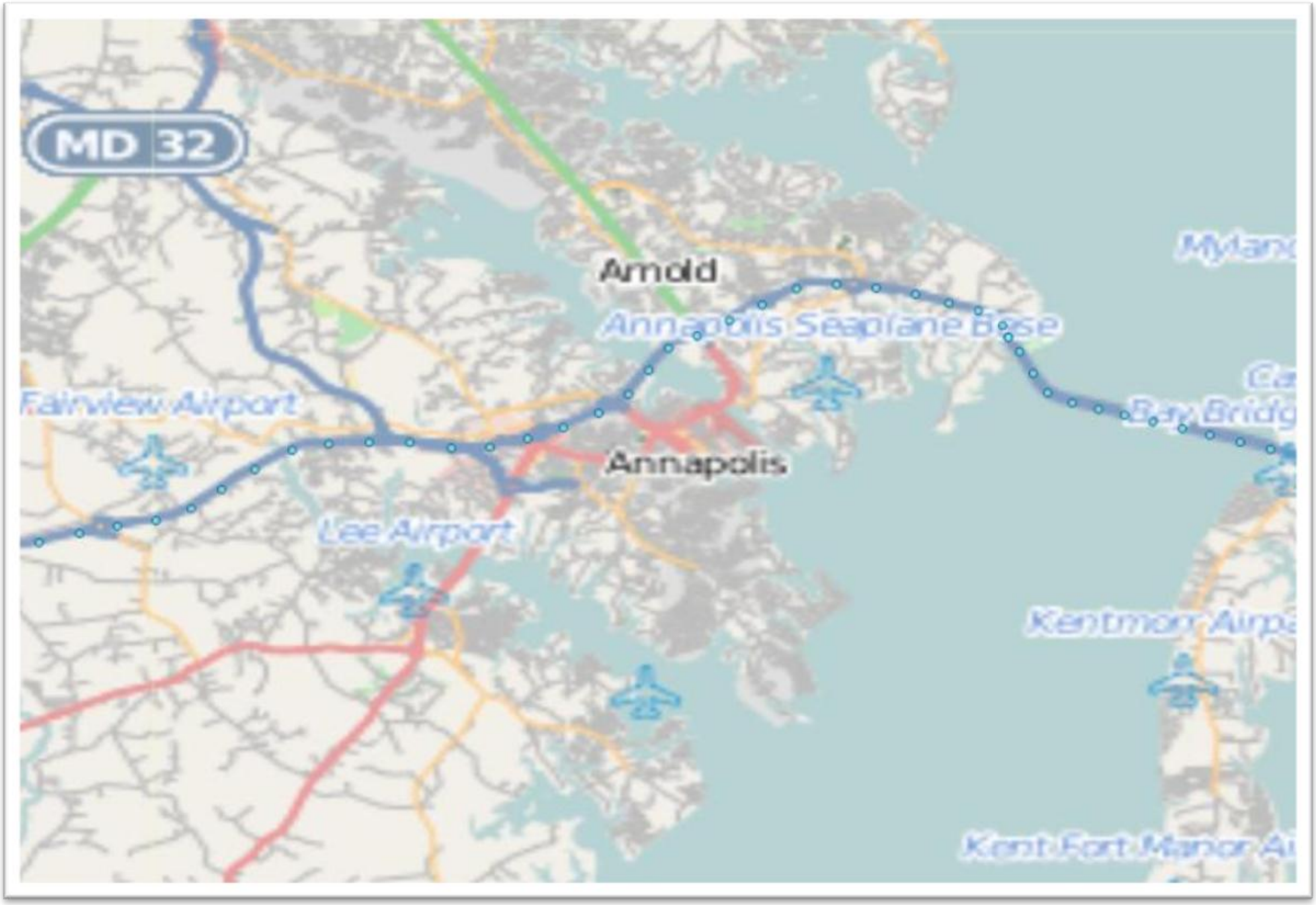


Figure 15: Zoom in of long distance trip

Once and individual dataset is checked for accuracy, it is saved to the operating computer and the server for backup. After all data is checked for accuracy, of the 263 original participants, a total of 218 usable datasets are available for analysis.

3.11 Data Analysis

3.11.1. Data Excluded from the Analysis

The GPS dataset is an integral component to this research topic. The reason why it has been difficult to determine the value of time using real word GPS data, is that the type of dataset necessary for this type of work has not been widely available.
The GPS devices were received in the first week of February. The devices were checked for robust data, and compared to the online trip diary to insure viability. After these steps were complete, a usable dataset of 218 participants was established. However, for this research, only participants who took the toll road during the survey period are usable in the calculation of value of time due to the methodology established. Of the 218 viable participants, 113 took the new toll road at some time in the 70 day trial period. Due to the fact that learning behavior is necessary for the value of time calculation, only participants who took the toll road for a pre-existing or redundant trip origin and destinations (only for trips that have occurred or will occur without the toll road) . After participants were removed that only took the toll road for unique trips, 58 individual's datasets were viable, and included in the value of time analysis.

3.11.2 GPS Data Processing

For all data processing efforts, the VBA code developed to take the raw data into the format used for this research's in depth analysis is included in the Appendix (section 8) This includes code to 1. Convert raw GPS points into a trip format giving basic variables on individual trips, 2. Creating a unique O/D and route identifier for each trip that is unique to the user, 3. Converting latitude longitude distances into route distances, 4. Calculate toll cost per trip via GIS roadway network.

The GPS dataset assigned a recording interval of 1 minute per data point. This interval has been used consistently in previous studies. The first step is to calculate the trip ends and hence origin destination points for each trip. A trip was determined as any series of points that had a velocity greater than 0 miles per hour, for more than three consecutive minutes. This definition of a trip is consistent with multiple previous studies (Wolf 2000). A number of trip statistics is calculated on the trip level, with travel time and distance amongst the most important for this study. This dataset of trips is loaded into GIS in order to get route and origin destination identifiers. After loading the US roadway network and creating a new shapefile and buffer for the toll road (ICC), a new variable for location of each point is created. Using this relative location to the toll road, a cost can be calculated dependent upon the time of day and number of points on the ICC. Using a code and the cost matrix provided by the Maryland State Highway Administration, a toll cost by trip can be calculated.

ICC Intercounty Connector — MARYLAND 200 — E-ZPass MARYLAND

E-ZPass Toll Rates Two-Axle Vehicles and Motorcycles

PEAK PERIOD Weekdays: *6 – 9 am, 4 – 7 pm (excluding federal holidays)*

FROM \ TO	I-370	Georgia Ave. (MD 97)	Layhill Rd. (MD 182)	New Hampshire Ave. (MD 650)	Columbia Pike (US 29)	Briggs Chaney Rd.*	I-95
I-370		$1.45	$2.00	$2.70	$3.35	N/A	$4.00
Georgia Ave. (MD 97)	$1.45		$0.60	$1.30	$1.95	N/A	$2.60
Layhill Rd. (MD 182)	$2.00	$0.60		$0.75	$1.35	N/A	$2.05
New Hampshire Ave. (MD 650)	$2.70	$1.30	$0.75		$0.65	N/A	$1.35
Columbia Pike (US 29)	$3.35	$1.95	$1.35	$0.65		N/A	$0.70
Briggs Chaney Rd.*	N/A	N/A	N/A	N/A	N/A		$0.70
I-95	$4.00	$2.60	$2.05	$1.35	$0.70	$0.70	

OFF-PEAK PERIOD Weekdays: *5 – 6 am, 9 am – 4 pm, 7 – 11 pm*
Weekends and federal holidays: *5 am – 11 pm*

FROM \ TO	I-370	Georgia Ave. (MD 97)	Layhill Rd. (MD 182)	New Hampshire Ave. (MD 650)	Columbia Pike (US 29)	Briggs Chaney Rd.*	I-95
I-370		$1.15	$1.60	$2.20	$2.70	N/A	$3.20
Georgia Ave. (MD 97)	$1.15		$0.50	$1.05	$1.55	N/A	$2.10
Layhill Rd. (MD 182)	$1.60	$0.50		$0.60	$1.10	N/A	$1.65
New Hampshire Ave. (MD 650)	$2.20	$1.05	$0.60		$0.55	N/A	$1.05
Columbia Pike (US 29)	$2.70	$1.55	$1.10	$0.55		N/A	$0.55
Briggs Chaney Rd.*	N/A	N/A	N/A	N/A	N/A		$0.55
I-95	$3.20	$2.10	$1.65	$1.05	$0.55	$0.55	

OVERNIGHT Every Day: *11 pm – 5 am*

FROM \ TO	I-370	Georgia Ave. (MD 97)	Layhill Rd. (MD 182)	New Hampshire Ave. (MD 650)	Columbia Pike (US 29)	Briggs Chaney Rd.*	I-95
I-370		$0.60	$0.80	$1.10	$1.35	N/A	$1.60
Georgia Ave. (MD 97)	$0.60		$0.40	$0.55	$0.80	N/A	$1.05
Layhill Rd. (MD 182)	$0.80	$0.40		$0.40	$0.55	N/A	$0.85
New Hampshire Ave. (MD 650)	$1.10	$0.55	$0.40		$0.40	N/A	$0.55
Columbia Pike (US 29)	$1.35	$0.80	$0.55	$0.40		N/A	$0.40
Briggs Chaney Rd.*	N/A	N/A	N/A	N/A	N/A		$0.40
I-95	$1.60	$1.05	$0.85	$0.55	$0.40	$0.40	

**Drivers can access eastbound ICC/MD 200 from Briggs Chaney Rd. Briggs Chaney Rd. may only be accessed from westbound ICC/MD 200.*

mdta.maryland.gov ezpassmd.com

Figure 16: Inter **County** Connector Toll Structure

3.12 Toll Cost

The GPS' logger recording interval is once per minute. Due to the fact that the shortest segment of the toll road is over two miles, it is not possible for the GPS logger to miss a segment and not record it. The GPS points were uploaded into a GIS environment and given a variable for each segment of the ICC that the individual points lie on. For each trip, these totals were added up and an accurate representation of the ICC route activity is achieved. Dependent upon which segments are used on the trip and time of day, the cost is looked up from the above cost matrix and allocated to the trip appropriately.

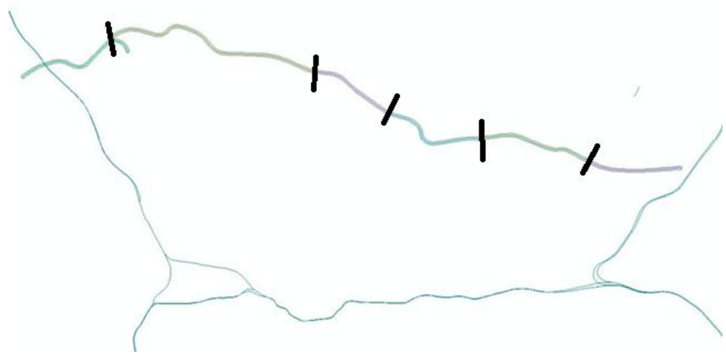


Figure 17: ICC roadway without GPS points in 6 segments

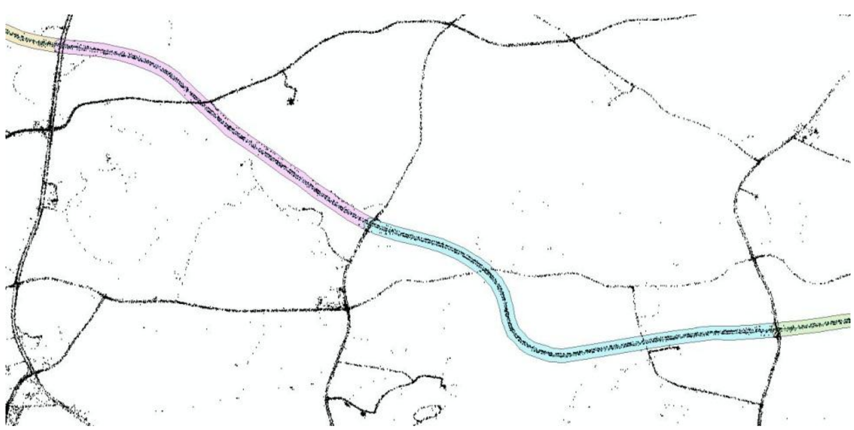


Figure 18: ICC roadway close up shown with GPS points

A separate code was written in order to create a unique origin destination identifier in order to group the trips with similar characteristics for modeling. The importance of this step is that a dataset with similar trip characteristics can be used in order to get a realistic change of behavior for the same trip over a period of time. To compare trips without the same origin and destination does not make logical sense, as a person's route availability (toll road availability) changes drastically.

The final dataset for this analysis includes 58 individual participants for 925 trips. This averages about 16 trips per individual. These are not necessarily the same origin/destination for all of a single individual's trips, but include only origin/destination trips that have the ICC toll road as a feasible option. In this way a dataset is created in which the user's travel patterns are uniform for all recorded trips, as well as being uniformly viable to a tolled trip; the bias for infeasibility of a route choice is essentially removed. The next section will show a step by step methodology for the calculation of daily value of time.

Chapter 4: Innovative Methodology for Value of Time Calculation

The objective of this module is to derive value of time based upon data collected in a real world environment. It is also imperative that this calculation of value of time be based upon real world information about the transportation network as well as the perceived costs and benefits from the interaction users have with their network. Keeping these objectives in mind, a methodology was developed in order to use real world data to create an adaptive learning system from the road network and individual driving behavior. Using GPS data, we look at how an individual travels for a normal origin destination pair, and infer what the user perceives about this trip. This information includes trip duration and cost. Keeping a catalog of back trips, individuals are able to make some assessment about their upcoming trip and what the anticipated time and cost will likely be. What this research does, is calculate these perceptions about their normal travel behavior, and calculate the individual's value of time dependent upon changes to their decision.

The way that this is done is by breaking down travel behavior into two categories: days in which an optional toll road was taken, and days in which an optional toll road was not taken. An individual user's perception of these two route types are recorded on a daily basis dependent upon previous trip information. Over time, the user has knowledge of both the toll and non-toll road, including time and cost. It can logically be deduced that, as an individual takes a toll road for a particular day, their value of time is the difference of perceived travel time for non-toll road to toll road travel time, divided by the cost difference between the two trips. This is the amount the user was willing to pay in order to reduce his travel time, based on the information that the user has at his/her disposal (previous knowledge). On days in which the user decided to take a toll road, the value estimated must be the minimum value of time for that day, due to the fact that the user was willing to pay at least that price, and it cannot be known whether or not they would have been willing to pay more. A day in which a user decided against taking the toll road, the value of time calculated must be a maximum, because they decided against paying the price in order to reduce travel time, and it is not known at what price they would have decided to reduce travel time for. Using many days' decisions, with constant updates to travel time and cost using real world GPS data, a picture of the average travel time for the user as well as variation in their day-to-day travel time can be developed.

4.1 Estimating Value of Time and Its Variation

This module in the positive behavior modeling approach is a separate portion that is calculated iteratively and based upon the knowledge that each individual possesses. While each individual's knowledge is positive (as opposed to normative), the value of time is calculated from their real world travel behavior and not based upon subjective beliefs.

For each individual, the toll road is a viable option for their normal travel. All participants who we have deemed toll roads unviable have been removed from the dataset. Take the next dataset as a short example of the calculation for a user's value of time for one origin/destination pair. The methodology will be explained step by step for this mock dataset:

Table 3: VOT Calculation Example

Day	Toll Road	Trip Duration (min.)	Toll Cost	Average Trip Duration (ICC)	Average Trip Duration (no ICC)	Value of Time Cap	Value of Time Floor
Day1	No	47	$0		47	N/A	N/A
Day2	No	52	$0		49.5	N/A	N/A
Day3	Yes	36	$3.20	36			$14.22
Day4	No	45	$0		48	$16.00	
Day5	Yes	35	$3.20	35.5			$15.36

- Day 1: For the first day of the origin/destination pair, the user is setting a standard for travel (by either taking the toll road, or taking the non-toll road). The average travel time for that route is recorded and the average trip duration is calculated. For the first day, the average trip will always be equal to the travel time for that day. For this day, the value of time cannot be calculated since there is no alternative route to base the change of behavior on.
- Day 2: The trip type is noted and the average trip duration is calculated once again. Since this trip is again taking place on the non-toll road, the trip duration is averaged with the previous day's trip duration. Again, the value of time cap cannot be calculated because there has yet to be a different route taken.

- Day 3: This is the first day in which the user has decided to take the alternate route. Similar to previous days, the average trip duration is calculated for that route. Since this is the first day in which this route is taken, the average trip duration will always be equal to the trip duration for that day. The calculation for value of time can now begin. For days in which the toll road is taken, the theory behind the Value of Time calculation is: The difference in money spent from toll to non-toll roads, versus the time savings for the individual trip. This is calculated by the difference in average time from toll to non-toll roads, multiplied by the cost for that individual trip. Calculation : (60/(49.5-36))*3.20 = $14.22/hour. This is the minimum value of time for that user for that day, since we know that they would be willing to at least pay this price, and possibly more.
- Day 4: The user now switches back to the non-toll road route option. The averages trip durations are updated for the non-toll road option. The value of time is calculated slightly differently than before. In this step the value of time is calculated as the time 'paid' in order to save monetary cost. The average trip duration for toll road is subtracted from the non-toll option (48 minutes-36 minutes) then multiplied by what the user would have had to pay in order to reduce their travel time ($3.20). The final value of time maximum for the day is thus (60/(48-36))*$3.2 = $16.00/hour. Contrary to the previous step, this is the maximum value of time for that user for that day, since they were unwilling to pay this value for the savings in time.

- Day 5: Once the average trip duration is updated for the toll road route option, the calculation method is identical to Day 3 using the updated average trip duration for non-toll road option from Day 4. Calculation: (60/(48-35.5))*3.20 = $15.36/hour.

As the value of time for the participant is calculated day by day, the overall value of time fluctuates to take into account the high value of time caps for days in which toll roads were used. Over a span of multiple days for multiple origin/destination pairs, these caps create a more trustworthy estimation for a singular value of time. The final value of time for each participant is calculated by averaging the value of time for each day in which it has been calculated (in this example, days 3, 4, and 5)

4.2 A note on possible negative value of times:

Using this methodology, it should be noted that many individuals would therefore have a supposed negative value of time. This can occur for a multitude of reasons.

1) The simplest case, for which this can occur, is that there was a small anomaly in the travel conditions for that day. A day in which a person decides to take the ICC may have a traffic accident along the route, increasing the travel time for the trip and thus creating a negative value of time for that trip. Since the previous day would most likely have a lower cost as well as travel time, and this user has decided on the other trip, this would calculate a negative value of time. Over the course of all trips for all OD's for the users, this small number of anomaly negative value of times should be countered by a much larger number of positive value of times.

2) The individual may take a chance on the ICC toll road to hopefully reduce travel time with true evidence or knowledge of a reduced travel time. In this calculation approach, this will result in a negative value of time for that day. Seeing as the rest of the individual's travel days are theoretically choice less (can not reduce travel time via toll road), the value of time remains negative. These trips are noted and removed from the calculation of value of time and averages for the remainder of the data.

3) The value of time can be calculated as negative using this approach while the toll road continues to be revisited, while still being logical from the point of view from the individual. These instances are hard to identify, as it is not clear as to why these choices are continually made. However, a situation can be feasibly understood if the main deciding factors for the individual and their route choice is not time or money. The individual could be deciding on trip routes dependent upon congestion for the day, or late departure from the origin location. Until more is known about each individual trip, and why some users decided to continue to pay a toll in order to increase their travel time, these explanations cannot be made. It may also be theoretically possible that some users have a negative value of time, but more reasonably, prefer to take the ICC toll road due to low congestion levels.

Below is an example of an individual who continues to take the toll road even though the cost is between $1.10 and $2.20 per trip and the average trip time is 5 minutes longer for ICC trips. This is a trip in which the individual takes

16 different times from the same origin to the same destination first taken on December 5th 2011, with the last trip occurring on February 2nd 2012.

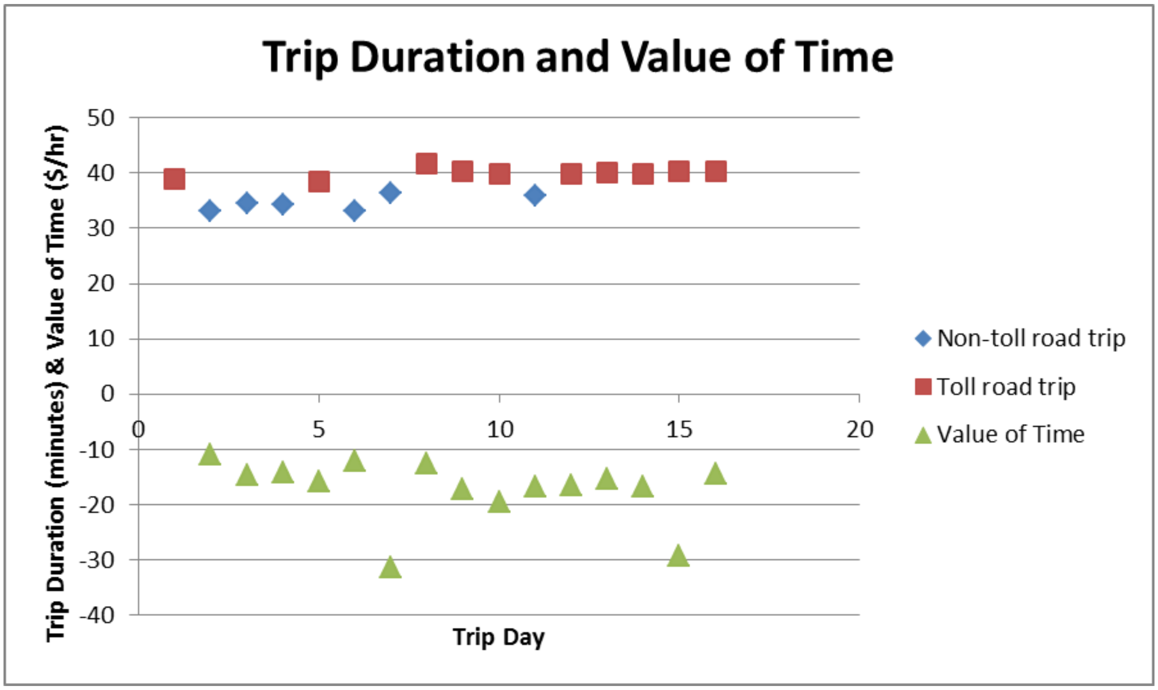


Figure 19: Negative Value of Time calculation

4.3 Results

The results section will cover VOT estimations in regards to both interpersonal and intrapersonal variation. The portion on intrapersonal variation will explore day-to-day changes of the user's value of time and how real world decisions impacted this change. The interpersonal results portion will aggregate all VOT estimations and show the distribution of these values across all participants for the entire survey period.

4.3.1 Day-to Day Variation of VOT within the Same Individual:

It is important to note the variation in day-to-day value of time, and in rare cases, small changes can be seen in the same day. Take the example below for a single participant (ID 39).

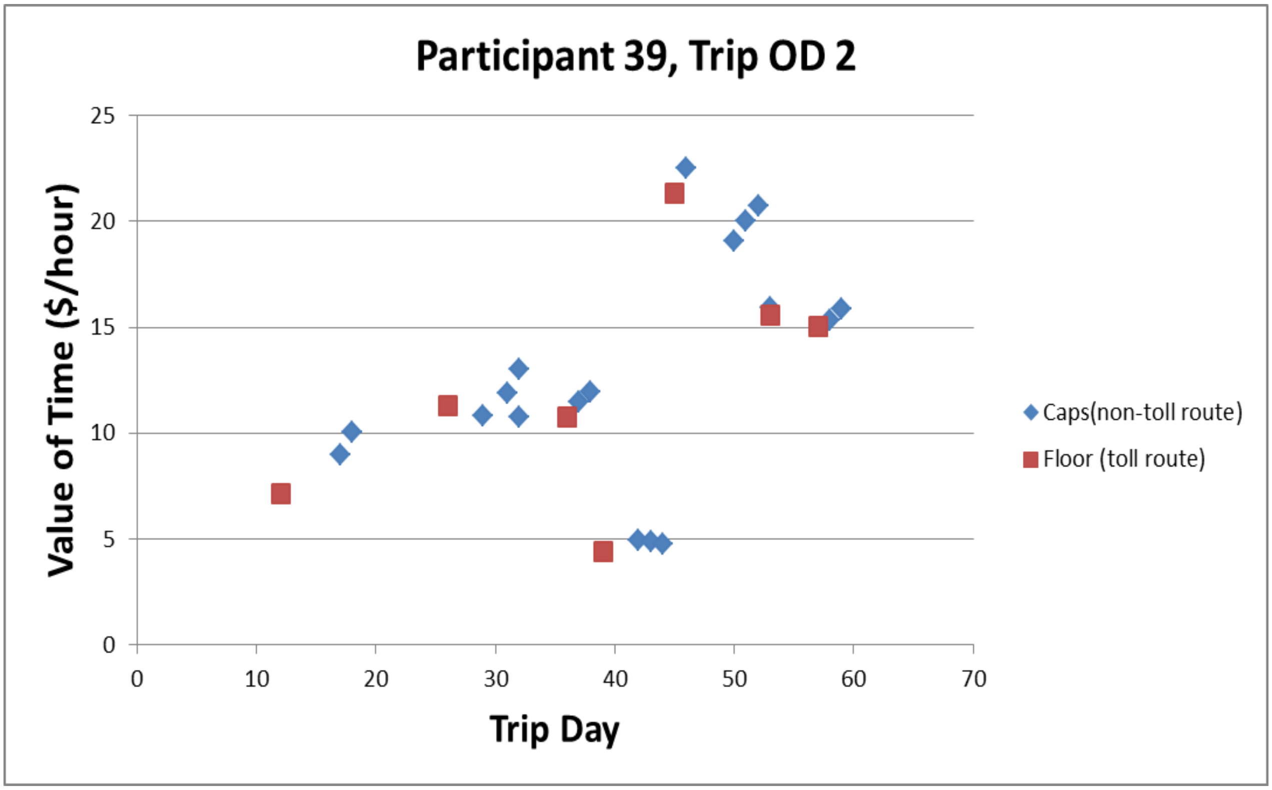


Figure 20: Individual participant value of time by day and trip OD

Looking at the day-to-day variation, it is interesting to view the results in this split manner. For this trip, the user's value of time is noticeably higher, with a rather uniform smattering of ICC trips. The variations of value of time are interesting to this individual, not only due to the uniform use of the ICC throughout the survey period, but the clear jumps in value of time. This occurs because the user decided upon not only the binary choice of whether or not to use the toll road, but also by what length of roadway to use, therefore changing his cost and time savings. The roadway is broken into 5 tolled sections (as shown in figure 2). On the first three toll days (days 12, 26 and 36), the user took 3 segments of the road for 7.8 miles at the cost of $1.55. However, we can see the drop-off in value of time on day 39, because the user decided to take only one section of the toll road for 2.24 miles at the cost of $0.50. The user then decides on another switch to their route choice on the toll road. On day 45, the user takes 4 segments of the ICC for 10.1 miles at a cost of $2.10. Finally, for the last two ICC trips, the user returns to the same route as the first 3 trips with his 3 segment, $1.55 cent approach.

It is a fair assumption to say that this is a normal commute trip for participant #39, and dependent upon personal situations, their value of time is different from day-to-day, willing to pay more for decreased travel time.

The valuable aspect of GPS data is that we are able to tell the exact travel pattern of an individual and visualize their decision making. Below, it is shown the progression of toll road routes the participant used in order to more accurately mold time and cost to their real world value of time.

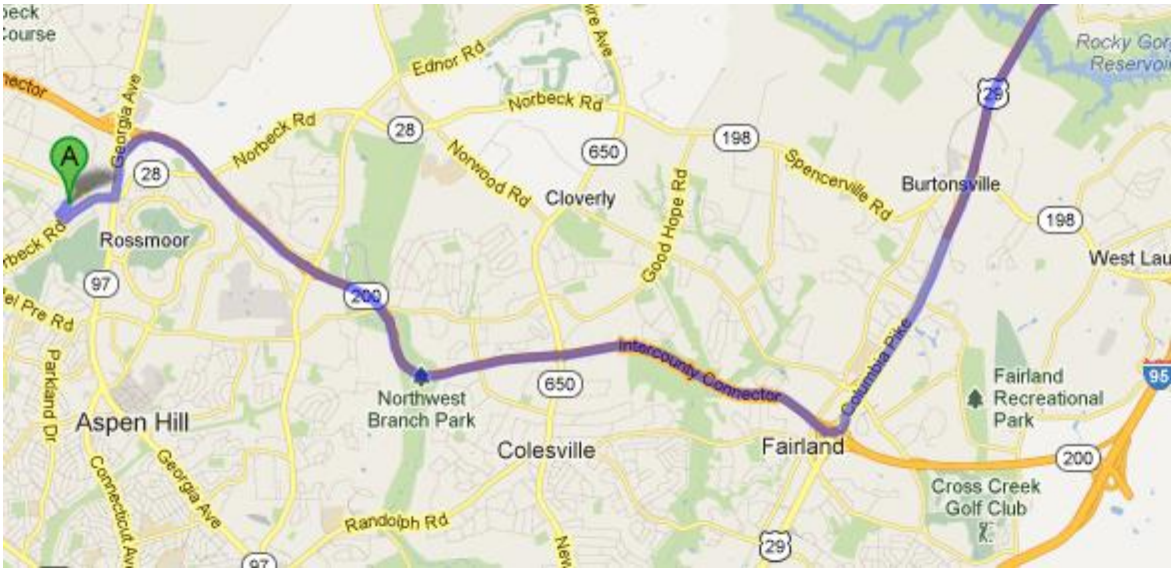


Figure 21: Days12,26,36,53,57: Three segemts, $1.55

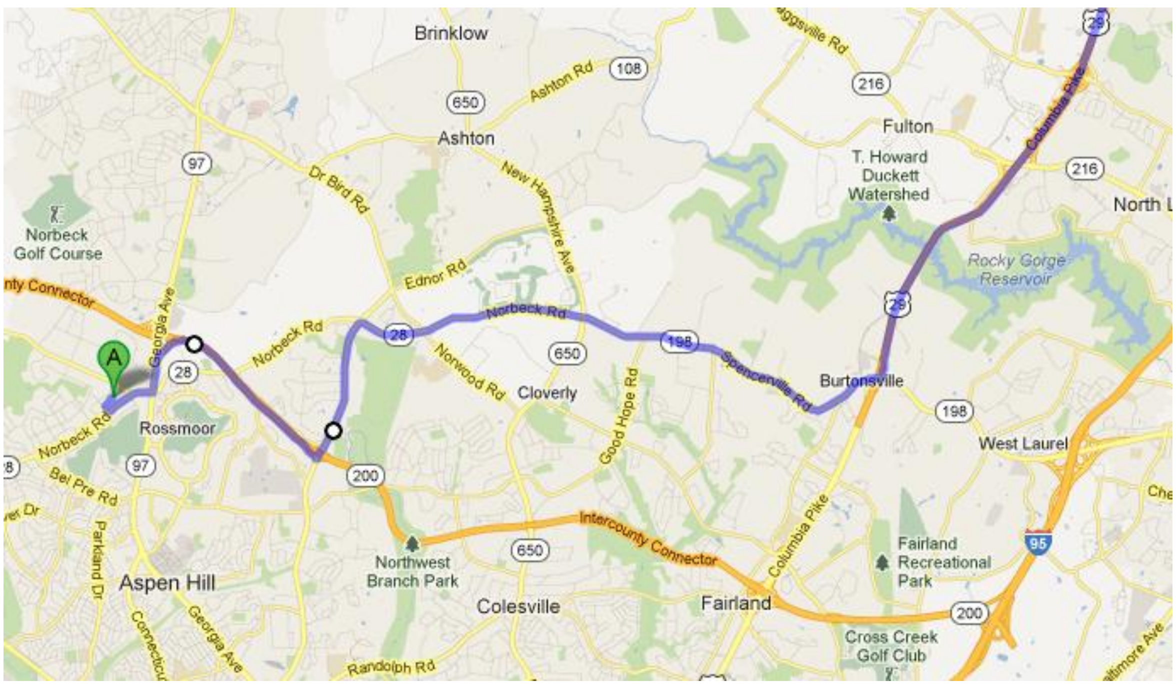


Figure 22: Day 39: One segment, $0.50

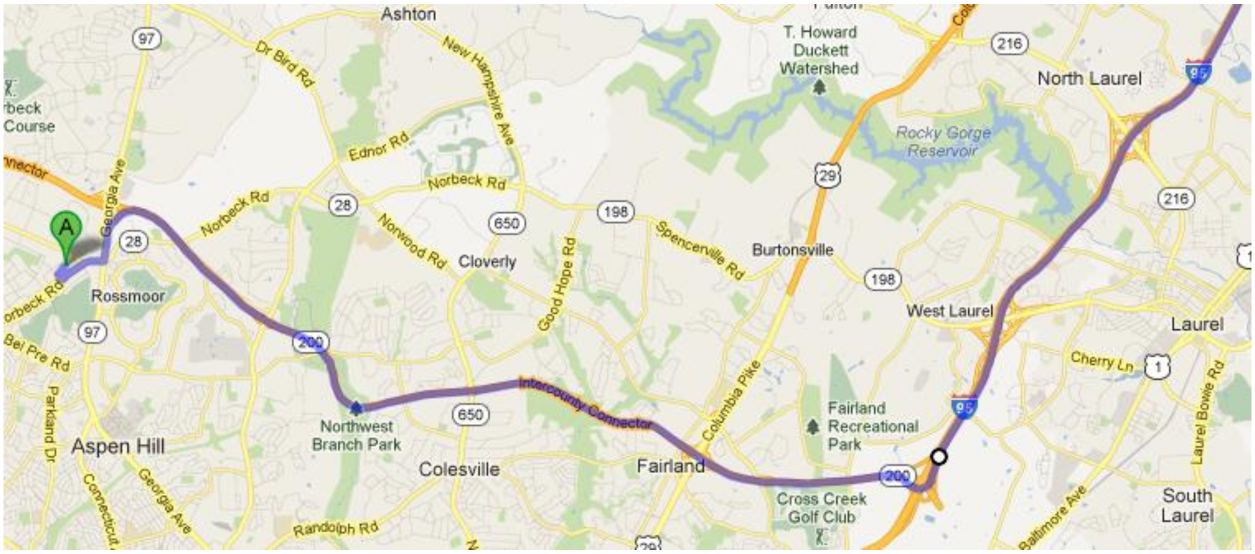


Figure 23: Day 45: Four segments, $2.10

This result helps to illustrate the complex choices individuals make with respect to value of time that cannot be evaluated using more traditional methods, that are possible with GPS based calculation of value of time.

4.3.2 Variation of VOT Across Individuals

The examples, methodology and results within the individual have shown VOT caps and floors on the day to day basis. In this section, results will be aggregated, and a single value of time will be given to each participant in the GPS survey. Under optimal roadway conditions, and perfectly rational behavior with no change in day-to-day VOT, these caps and floors will result in a regimentally narrowing window in which the individual user's value of time will lie. However, due to the possibility of roadway congestion, changes in user preferences, and logical behavior, it is possible that a cap for VOT is lower than a floor on a different day. The way the results are handled therefor, is that for each value of time estimation shown in this section, it will be calculated using the average of all floors and caps for the individual.

Results show a wide range of individual values of time. They range from a near complete unwillingness to pay for travel time reduction at $0.34/hour, up to 27.3$/hour, for the highest value of time in the survey. The reason why these results are intriguing is due to the nature of the methodology; the use of real world longitudinal GPS data gives a level of accuracy to value of time estimation that could not previously be achieved. The graph below shows the value of time per individual.

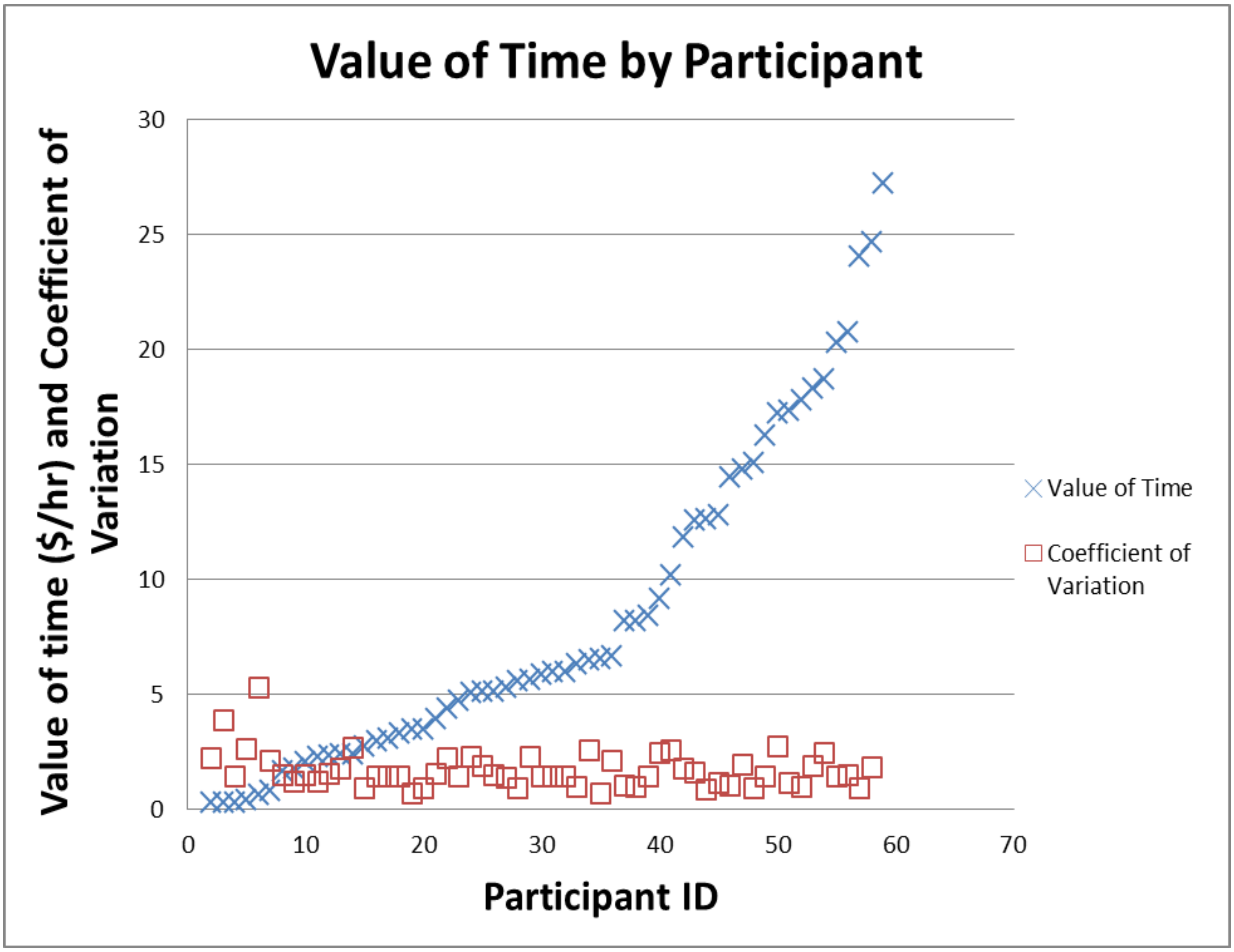


Figure 24: Full dataset Value of Time averages vs. their coefficient of variation

Included with the average value of time over all trips is the coefficient of variation amongst those trips. This is quantified as the standard deviation divided by the mean of all trips. IT is evident that those individuals with the lowest value of time tend to have the highest coefficient of variation. This is most likely due to the fact that they rarely take a trip on the toll road, having a usually low value of time, but in an instance when the toll road is taken, the variation is increased greatly as compared to the mean.

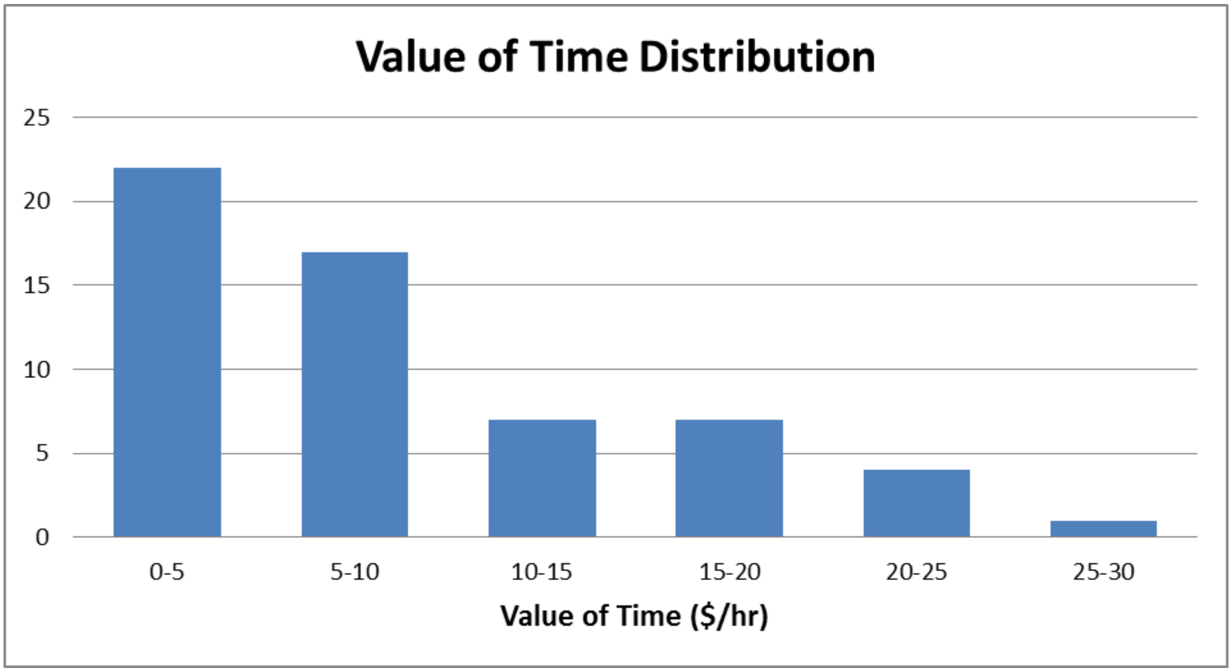


Figure 25: Value of Time Frequency

The distribution of value of time shows the low value for most participants in the survey. The largest group is willing to pay between zero and five dollars for a reduction in one hour of travel time. Each subsequent price range reduction accounts for a smaller amount of individuals. The average value of time for all individuals is $8.34 per hour. This is a rather low value of time when comparing it to previous papers studying value of time with different methodologies. The reasoning for the low value of time estimations can be due to the fact that these are not necessarily commute trips, making the VOT lower, but this would be purely speculative. While these values are low, they are based upon real world decisions made by users and accurately reflect the user's true value of time.

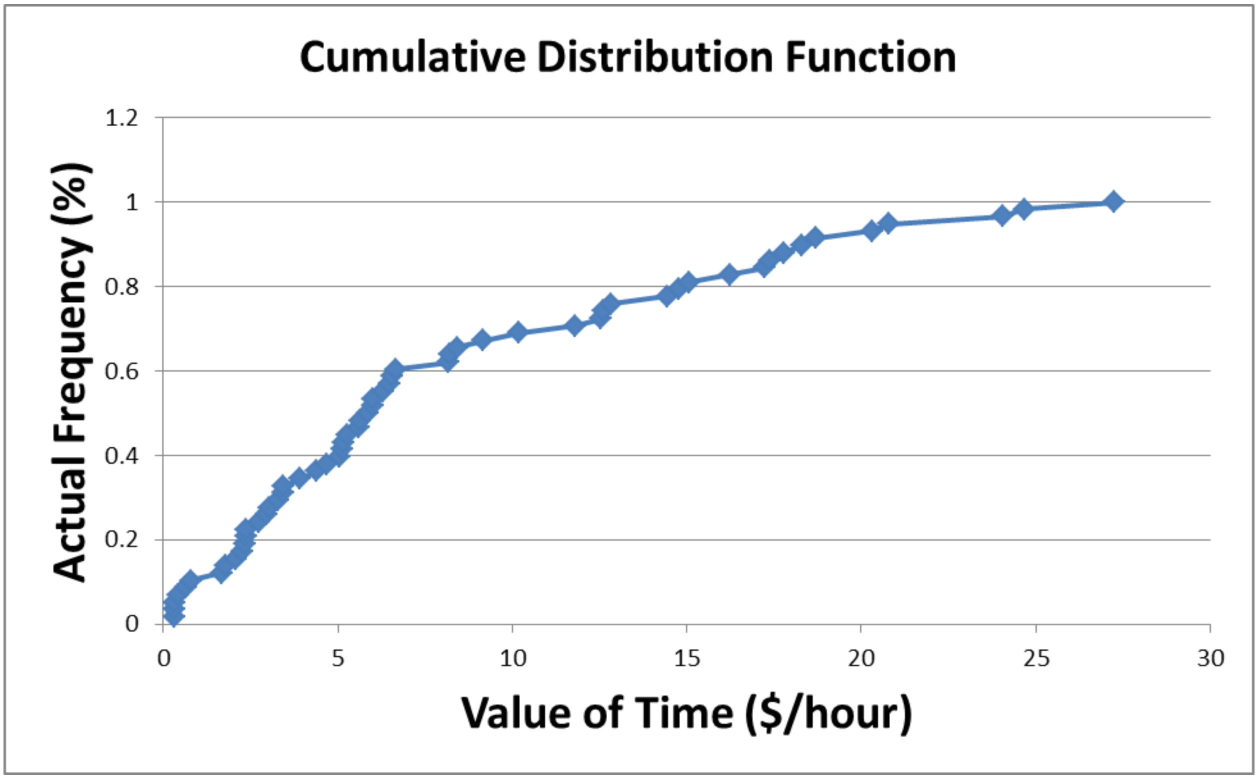


Figure 26: Value of Time Cumulative Distribution Function

The cumulative distribution function shows similar results to the probability distribution previously. The majority of values (% 60) lie below the $7/hour mark. 20% of all participants' value of time averaged over $15/hour.

4.4 Comparison with previous VOT estimates

Keeping in mind the major difference in approach for calculating value of time between this study and other studies that use mixed logit models, it is important to study the validity of the results provided in this work. Three different models are used to show the fluctuating estimations for value of time. Reasoning for the difference in value of time may very well be the difference in datasets used in the different models. For example: Small et al.'s (2005) model primarily dealt with commute trips, whereas Cirillo and Axhausen (2006), Hess et al.(2005) and Zhang and Krause (2012) don't necessarily look at commute trips. Of course, the present study is the only one which implements real world GPS data, whereas others deal with RP, SP, or simulated data.

Table 4: Comparison of Value of Time Estimations

Modeler	Model	Data	Value of Time
Small et al. 2005	Mixed Logit	RP/SP Data	21.46$
Hess, Bierlaire, Polak 2005	Mixed Logit	Simulated Data	Shows wide range in VTTS estimation
Cirillo and Axhausen, 2006	Mode Choice, Mixed logit	SP Data	12$
Present Study: Krause, 2012	VOT Cap procedure	GPS Longitudinal	$8.34

It should also be noted, that this calculation does not strictly consider commute trips. The theoretical value of time for an individual taking a commute trip should be higher than that of an individual who is not necessarily taking a commute trip. Small et al (2005) used a dataset primarily for commute trips.

Chapter 5: Determination of Heuristics for Route Choice Behavior

The calculation of Value of Time has been thoroughly explained using the updating travel experience of individuals. While the theoretical model has been explored, each module in the positive behavior model has not been studied in depth. In this section the two modules for explaining the user's search and route decision making behavior will be analyzed. The If-Then rule set for each will show the series of heuristics that each individual goes through with their knowledge in order to make a decision. The data set used is the same used for calculation in the VOT module. All updates made from the VOT module are kept in the knowledge base, so the VOT estimation can be used in the heuristics for determining search and decision rules. Keep in mind that only trips that have already occurred are included in the individual's knowledge base, and can therefore be used for the days' search and decision rule heuristics. Information about upcoming trips cannot be taken into account.

The heuristics were determined using the WEKA machine learning software. The JRip algorithm was used to select rules and trim the total rule set to a manageable and optimal level. The full list of variables included in both rule sets as well as the classification method and confusion matrices are shown in Appendix D (page 89).

5.1 Search Rules

First, a classification of what a search implies must be described. If the individual does not change routes for the entire duration of the travel survey, then one can assume that all searching for alternative routes has stopped. If for instance, an individual takes the same commute trip for 22 days during the trial period and makes 5 switches in the first 15 days, but thereafter does not make any switches in route choice, then it can be said that that individual is no longer searching for an alternative route. For a trip to classify as including a search, some future trip for the same origin destination pair must be of a different route than the current route identifier. The only exception to this rule is if the individual makes regular switches to their route up until the ending of the travel survey. Since the survey ends at a relatively arbitrary time, it would not be correct to deem any trip at or near the end of the survey as not searching (since there would be no later trips with a differing route). Instead, any trip that is within the last 4 trips of the end of the survey, and the user was otherwise searching for alternative trips up until that point, those trips are also deemed searching trips. Reverting back to the theoretical framework, it can be followed to show the impact of the searching module. If the user decides to search for an alternative route, then the next module (decision rules) will occur. If the user decides against searching, then repetitive behavior takes over, travel is experienced, information is learned, and knowledge is updated for the next iteration.

These rule sets explain 89% of search versus non-search behavior. Using previously calculated VOT estimations from the VOT calculation Module, the user's knowledge base, as well as demographic variables is used. Below is an explanation of all variables included in the rule set followed by the rules themselves, and an in depth explanation of each is finally given.

Variables explained:

- d_VOT: this is the difference between calculated value of time from the previous two days of the individual as calculated by the Value of Time module explained in previous chapters. When viewing a negative number, it does not necessarily mean that the user has a lower value of time, but rather that their value of time has decreased over the previous 2 trips. For example if on day 1 their calculated VOT is 25$/hour and day 2 the VOT is calculated at 10$/hour then the d_VOT for all days after day 2 before another trip on that route is taken will be -15.
- Distance on ICC: the distance, in miles, required to drive on the ICC for that selected route.
- Value of time: The most recent estimation of value of time for the participant prior to that travel day as calculated by the Value of time calculation module.
- Yesterday trip: 0 if trip was normal, 1 if trip was ICC, route for previous travel day.
- No_ICC_TIME: The average time for all trips previously traveled on the same OD that were not toll road trips.
- Sex: Gender of the participant.
- Average time combined: The average time for all trips previously traveled for the same OD. (Includes ICC and non-ICC trips).

Below are the search rules. If any one of these rules is true for the individual trip, then the user will not search for an alternative route:

1. d_VOT >= -16.5 and distance_on_icc <= 7.72 and value_of_time_true <=5.5 and yesterday_trip = no ICC

2. -6.1<= d_VOT <= 4.6 and Age >= 58

3. ICC_trip_distance_percent <= 5% and no_icc_time <= 64.1

4. Sex = Female and 1.5>=d_VOT >= -1.2 and distance_on_icc >= 10.1

5. 20.3=< average_time_combined <= 22.4 and Age <= 31

Then stop searching for alternative route. Else If; continue search.

Rule Set Explanation:

1. If the user has a low value of time (less than 5.5$/hour) and the value of time estimation has been decreasing (d_VOT>=-16.5) and the previous day was driven on the non-toll road route, if the distance is low on the ICC, searching will cease. (Correct 112 out of 128 times).

2. If the change in value of time is relatively stable (between 4.6 and -6.1$/hour change from the last two days) and the individual is greater than 58 years old, the searching ceases. (Correct 35 out of 41 times).

3. If the trip distance is less than 5% of the total trip length and the travel time for the non-ICC option is less than 64 minutes, then cease searching. (Correct 24 of 24 times).

4. Like other similar rules, if the VOT change stabilizes (-1.2$/hour to +1.5$/hour) and the distance on ICC is high, and the participant is female, then searching ceases. (Correct 13 of 14 times).

5. If the user is below 31 years old, and the average trip time has stabilized (20.8 - 22.5) then searching will cease. (Correct 12 of 13 times).

Else if; continue searching. (Correct 728 of 778 times).

The overarching theme for a majority of these rules shows that once the user's value of time has stabilized (rules 1, 2, &4) that the user will stop searching. Also, when the average time for the route is roughly 20 minutes, then the user is satisfied with the route chosen, and tends to stop searching as well. As for the rule that is slightly different from the others (rule 3), it shows that if the toll road portion of the trip is a very low percent of the trip, the users tend to believe that it is not worth taking if the trip is long. This set of rules show that it is feasible to estimate when the user is open to alternative routes at a relatively high level of accuracy.

Now that there is some reasoning behind the searching behavior of users, we will move on to the route decision rules to see how they differ from searching.

5.2 Route Decision Rules

Once an individual has decided to search for the alternative route, they must determine whether or not they will switch to the alternative route. If they decide to switch routes, based upon knowledge thus far accumulated, then travel is experienced and information is learned and knowledge is updated. The way that it is calculated was previously shown, with the constantly updating average of time and cost for the individual O/D and route.

The rules shown below are based upon both socio-economic and trip data from previous days. In the same manner that the value of time calculations are based upon the knowledge database of the user, these rules are based upon the same knowledge base. However, when calculating value of time, the only factors that are necessary are the time and cost per trip. When using the knowledge base for decision and search rules, more information is used for each trip, such as speed and distance. Also, while not in the knowledge base, or rather, not intrinsic to their knowledge base is socioeconomic variables which are also used in the rule set. To be considered are all variables collected in the initial participation online survey:

- Age
- Income
- Gender
- Education

This does not necessarily mean that these variables are significant for the individual travel behavior and are not used to determine route decision rules.

There may also be some impact of time of day on switching to opposite routes. The travel times were broken down into three segments: Peak, Off-Peak, and Night time. For instance, if a user usually takes a normal route which is a non-toll facility, and the current day's travel is during the peak time, it may be less likely for the individual to switch to a toll road during the peak time due to the increased cost of the toll facility. It may also be more likely that an individual with high income is more likely to use the toll road during peak time because the relative savings in travel time for cost in a congested area during peak congestion time. These are some intuitive impacts of recorded variables in the GPS data, however, the rules derived using the JRip learning algorithm do not deem time and income as important factors for route choice. It is important to note however, that these variables are considered, but simply are not significant in the user's choice. The variables that had the most explanatory power can be viewed in the rule set below. For a full set of considered variables, please visit the raw model outputs section of the Appendix (D).

Rules: These rules explain 80.2% of the route decisions by participants in the 70 day survey. Based on information learned through previous day's travel, these rules were determined to be the most explanatory in deciding upon trip route. The learning program JRip was used to determine the heuristics based upon its' pruning function to limit the rules necessary to explain the largest amount of behavior. Variables in this rules set include:

- Age (age of the participant in years),
- ICC_savings (this is calculated as the amount of time the user believes they will save by taking the ICC. This is not necessarily based upon real world travel experience if the user has yet to take the ICC for that route. If the ICC has not be taken on that route previously, the time savings is calculated by calculating the new time based on a free flow speed on the stretch of roadway that will be used for the ICC added to the average speed for the remainder of the trip),
- ICC% (the percent of the trip that will need to be taken on the ICC in order to arrive at the destination. For example, a trip that is 10 miles in total, and 5 miles are on the ICC will have an ICC% value of .5),

- Distance_ICC (the distance traveled on the ICC),
- Distance (average distance for that OD for that user),
- Yesterday Speed (the average speed of the user's trip from the previous day).

Rules for Choosing ICC route:

1. If Age <= 37 & ICC_savings >12 minutes & ICC%<.85

2. Or Distance_ICC >7.8 miles & Age <= 42 & .53<ICC%<.67

3. Or Distance>31 miles & ICC%>.24

4. Yesterday Speed< 10 miles/hour & ICC_savings<-14 minutes

5. ICC_savings>31 minutes & Distance>8.6 miles & Cost>=$0.8

Then choose ICC. Else if; choose non-ICC route.

Each rule is explained in depth:

1. If the user's age is below 37 years, and they believe that they can save 12 minutes on their route by taking the ICC and the ICC portion of the trip will not take more than 85% of the total trip length, then they decide on the ICC for that trip. (Correct 49 of 52 times). This seems to show that relatively younger participants, when their experience says that there will be a time savings with ICC, are highly likely to take the ICC route.
2. If the distance of the trip requires an ICC portion longer than 7.8 miles, the user is less than 43 years old, and the ICC portion makes up between 53 and 67% of the total trip length, then the ICC is taken. (Correct 18 out of 18 times). Again this shows that if the distance is reasonable, and the percent of

the trip taken on the ICC is right, then the toll road will be taken for relatively younger individuals.

3. If the trip length is greater than 31 miles and the ICC will take at least 24% of the trip length, then take the ICC. (Correct 30 out of 36 times). This is a rule not dependent upon learned behavior, (the only one of the rule set), which basically says that if the route is over a certain distance, and at least 24% of that trip can be on the ICC, then the ICC will be used.
4. If yesterday's trip has an average speed of less than 10 miles per hour and the user believes that they will spend more time on the ICC by 14 minutes than their alternative route, they will take the ICC. (Correct 7 out of 7 times). This is an extremely interesting rule derived from a relatively low number of selected routes (7). IT shows that learned behavior is very important to their selection of routes for that day. Even with an assumed loss of 14 minutes on their trip, participants were willing to take the ICC route to increase their average speed. To some individuals, it shows that their travel speed is more important to their travel time.
5. If the user believes they can save 31 or more minutes on a trip whose distance is greater than 8.6, and involves the segments to the west of route 650, or includes more than one route (due to the toll cost), then the ICC trip is taken. (Correct 15 of 16 times). This is the only rule that takes into account toll cost (as a form of segment location), which shows that rather than wanting to limit cost, cost is mostly a function of route desirability, and will only take the toll road if the location of the route takes the correct segments. This could show

that the routes to the east of 650 have a lower desirability for users who are attempting to save travel time while maximizing distance.

Else If; Take non-ICC route option. (Correct 805 of 949 times).

It should be noted that the single most explanatory variable when choosing between toll and non-toll road routes, age is the most important factor.

The variables with the highest impact on route decision (deciding between toll and non-toll road) ranked from most important to least are:

1. Age
2. Distance traveled on ICC
3. Total Trip Distance
4. Cost of ICC toll
5. Time of trip on ICC
6. Change in value of time from previous days
7. Average speed of trip off of ICC
8. Change in time of trips compared to previous trips
9. Is the first trip of the GPS survey for that O/D pair

Age is a remarkably good identifier however, almost surprisingly. For participants under the age of 35, 63% of all trips in the dataset are taken on the ICC. Over the age of 35, 24% of trips are taken on the ICC. The remainder of the variables with the greatest impact on route selection tend to make more intuitive sense, with distance traveled on ICC being next highest, followed by Total Trip Distance and Cost of the toll.

Using this rule set it is possible to determine the likelihood that an individual will take the ICC toll road. As an application, the rules can be useful in policy analysis in order to optimize toll road ridership levels. It is also important to understand why users take the toll road and these rules could even be used in the future to design a new toll road system.

The next section will conclude the thesis by reviewing the research contributions accomplished and explore further applications of this research work for the future.

Chapter 6: Conclusions

The value of time calculation module has taken an in depth look at the way in which value of time has been calculated previously. In so doing, current methods of discrete choice analysis were compared to the new approach of using real world longitudinal GPS survey data to calculate individual value of time dependent upon day to day route choice. The GPS dataset was explored and how this new dataset is uniquely built for the work at hand. Using a method in which the value of time is incrementally capped and averaged based upon route choice, an estimation of the individual user's value of time is calculated. Based on information from previous days' travel, and the decisions made regarding cost and route for the current day, a reliable value of time estimation can be calculated for a span of several days and weeks. With the ability to gain insight on learned behavior and view their impacts of decisions made day–to-day, intricacies can be viewed that were previously hidden. This is evident when viewing the individual's data that changed the length and cost of toll road to take based upon cost and time preferences for that day. The unique combination of very accurate GPS data with trip learning behavior has led to a new, promising dimension of value of time calculation. The aggregate data also shows merit when compared to more traditional VOT calculation methods. Due to the large variability of value of time between participants, it can be said that a strict value of time calculation or distribution for multiple individuals in the realm of logit models may not be realistic in real world environments.

Some future work also has promising application in this field of real world VOT calculation. Using demographic information, as well as more in depth trip characteristics, it may be possible to explain why the user's value of time is changing. Information such as this can be a valuable addition to current and future policy decisions, especially in implementing toll cost structures for new toll roads.

Moving on to the theoretical framework of iterative learning, we can see that it is an important step in understanding the process that an individual takes in determining their route decisions based upon price and time for trips. While the accuracy of the rules is a rather low 80%, this is expected since the rules are based upon over 50 individuals' behaviors that are anything but uniform. If a longer survey was performed that looked at similar data for a singular participant, a set of rules can be derived that has a higher level of accuracy than those derived here. Many aspects of the model can be seen as important advances. The variables that continue to play a major part in the rules shown are that of time, cost, value of time, and duration of trip. Clearly, these are important factors in which people decide on route from previous day's travel experience. If, for example, the route taken on the previous day was the non-toll road route, and the travel time was greater than the previous days' average travel time, the likelihood that the individual will switch to the toll road route is greater. Also, surprisingly, the only socio-demographic that was ultimately included in the rule set was that of Age and Gender. Income and Education did not impact the individual's decision making process on the large scale. This may be largely due to the fact that those that frequented the ICC route and therefor had it as a viable option already have a high income that does not have much variation. This is the same with high education levels. This can also be seen in the Appendix which shows the breakdown of Education and Income as compared to the surrounding area. This is an inherent bias in the selection of individuals and could have caused a bias in the rule set.

This thesis has looked at a multistep and varying approach to understand individual's value of time using the first only GPS longitudinal dataset to do so. First, an in-depth look at the collection process for the dataset was shown. The methodology for data acquisition ranged from website and survey design, participant selection, data acquisition, and final data processing. Through the process of Value of Time capping procedures, a value of time estimation is derived for all individuals who have the toll road as a viable option for travel. Using the individual's travel patterns as a choice system for their value of time, estimations are derived day-to-day dependent upon their choice made. Furthering this work, the iterative learning framework is used as a means to better understand the information gained by the user from their day-to-day travel. Using the choices that the individual has made, heuristics were derived for their searching and deciding on their routes for normal travel dependent upon trip duration and cost. Learning about individual's choices and how they are impacted by cost and time has helped to further study the true value of time by individual. Using the approaches derived from this thesis, the derivation of value of time becomes a more solid basis with the use of real world longitudinal GPS data.

Using this approach, we are now not only able to calculate a real value of time for each individual on a daily basis, but also have a better understanding of why these decisions are made with the positive behavior approach. For the first time, theories can be adopted as to why a negative value of time can seemingly occur by viewing rules that emphasize other importances over time and cost (such as speed, and comfort, familiarity, etc…). This research has continued to build upon a pre-existing framework for travel behavior analysis that now uses real world GPS data to explain travel decisions in the positive behavior approach.

Appendices

Appendix A: Data processing code (VBA)

Raw data to trip format:

```
Sub check()

Dim original As Worksheet

Dim a As Long

Dim b As Integer

Dim OD As Worksheet

Set original = ActiveWorkbook.Sheets("Input")

Set OD = ActiveWorkbook.Sheets("Output")

a = 3

b = 2

While a < 50000

  ' If original.Cells(i, 2).Value - 0.001 > original.Cells(i + 1, 2).Value Or original.Cells(i, 2).Value + 0.001 < original.Cells(i + 1, 2).Value Or original.Cells(i, 4).Value - 0.001 > original.Cells(i + 1, 4).Value Or original.Cells(i, 4).Value + 0.001 < original.Cells(i + 1, 4).Value Then

    If original.Cells(a, 8).Value > 1 Or original.Cells(a - 1, 8).Value > 1 Or original.Cells(a - 2, 8).Value > 1 Then

   OD.Cells(b, 2).Value = original.Cells(a, 2).Value

   OD.Cells(b, 3).Value = original.Cells(a, 4).Value

   OD.Cells(b, 4).Value = original.Cells(a, 6).Value

   OD.Cells(b, 5).Value = original.Cells(a, 7).Value

   OD.Cells(b, 6).Value = original.Cells(a, 8).Value

   OD.Cells(b, 7).Value = original.Cells(a, 9).Value

   OD.Cells(b, 1).Value = original.Cells(a, 13).Value
```

```
   b = b + 1
   End If
   a = a + 1
Wend
'Dim original As Worksheet
Dim i As Integer
Dim j As Long
Dim k As Integer
Dim m As Double
Dim n As Double
Dim p As Integer
Dim q As Integer
Dim Z As Integer
'Dim OD As Worksheet
'Set original = ActiveWorkbook.Sheets("Input")
'Set OD = ActiveWorkbook.Sheets("Output")
Set Var = ActiveWorkbook.Sheets("Set")
lLoop = 1
i = 1
'iteration
j = 3
'row
k = 0
'#turns
m = 0
'distance
n = o
```

```
'summed distance
p = 0
'number of gps points per trip
q = 0
'number of stops on trip
Z = 2
'sheet 3 counter
While i < 5000
        If Abs(OD.Cells(j, 7).Value - OD.Cells(j - 1, 7).Value) > 90 And Abs(OD.Cells(j, 7).Value
- OD.Cells(j - 1, 7).Value) < 270 Then
        k = k + 1
        'check heading at each data point, if greater than 90 degrees, count turn integer'
        End If
  n = n + m
  m = Sqr((OD.Cells(j, 2).Value - OD.Cells(j - 1, 2).Value) ^ 2 + (OD.Cells(j, 3).Value - OD.Cells(j - 1,
3).Value) ^ 2)
  'calculates algebraic distance based upon lat long data'
  OD.Cells(j - 1, 9).Value = m
  'point to point distance'
      If OD.Cells(j - 1, 6).Value < 1 Then
      q = q + 1
      End If
  p = p + 1
  'iterates time counter'
     If Abs(OD.Cells(j, 4).Value - OD.Cells(j - 1, 4).Value) > 200 And OD.Cells(j - 1, 17).Value <>
59 Then
     'if the vehicle stops for more than 2 minutes then...'
```

```
    Var.Cells(Z, 2).Value = OD.Cells(j - p, 2)
    Var.Cells(Z, 3).Value = OD.Cells(j - p, 3)
    Var.Cells(Z, 4).Value = OD.Cells(j - 1, 2)
    Var.Cells(Z, 5).Value = OD.Cells(j - 1, 3)
 OD.Cells(j - 1, 8).Value = "Stop"
 'Destination of trip
 OD.Cells(j - 1, 14).Value = p - 1
 'inputs travel time for that trip
 OD.Cells(j, 8).Value = Abs(OD.Cells(j, 4).Value - OD.Cells(j - 1, 4).Value)
'Soak time from the end of one trip to the start of the other'
OD.Cells(j - 1, 12).Value = k
'number of turns over 90 degrees taken on this trip
OD.Cells(j - 1, 10).Value = n
'prints summed distances by trip'
OD.Cells(j - 1, 13).Value = q
'prints summed stops by trip'
'Var.Cells(Z, 9).Value = OD.Cells(j - 1, 4).Value - OD.Cells(j - p, 4).Value
 'inputs travel time for that trip
'Var.Cells(Z, 6).Value = k
'number of turns over 90 degrees taken on this trip
' Var.Cells(Z, 7).Value = n
'prints summed distances by trip'
'Var.Cells(Z, 8).Value = q
'prints summed stops by trip'
    lLoop = lLoop + 1
k = 0
n = 0
```

```
            m = 0

            p = 0

            q = 0

            'resets counters

            Z = Z + 1

            End If

  j = j + 1

  i = i + 1

  'iterates counter and row'

Wend

End Sub
```

Create unique OD identifiers

```
Sub Trip()

    Dim i As Long
    Dim j As Long
    Dim r As Range
    Dim T As Worksheet
    Dim C As Worksheet
    Set C = ActiveWorkbook.Sheets("Cost")
    Set T = ActiveWorkbook.Sheets("Trip")

   ' Dim Olat As String
   ' Dim Olong As String
   ' Dim DLat As String
   ' Dim DLong As String

    i = 1
    j = 1

    Set r = Range("B2:E122")
    While i < 300
        If C.Cells(i, 27).Value < 1 Then
            While j < 300
                If Abs(r.Cells(i, 1) - r.Cells(i + j, 1)) < 0.003 And Abs(r.Cells(i, 2) - r.Cells(i + j, 2)) < 0.003
And Abs(r.Cells(i, 3) - r.Cells(i + j, 3)) < 0.003 And Abs(r.Cells(i, 4) - r.Cells(i + j, 4)) < 0.003 Then
                 C.Cells(i + j + 1, 27).Value = n
                 C.Cells(i + 1, 27).Value = n
                End If
```

```
            If Abs(r.Cells(i, 1) - r.Cells(i + j, 3)) < 0.003 And Abs(r.Cells(i, 2) - r.Cells(i + j, 4)) < 0.003
And Abs(r.Cells(i, 3) - r.Cells(i + j, 1)) < 0.003 And Abs(r.Cells(i, 4) - r.Cells(i + j, 2)) < 0.003 Then
            C.Cells(i + 1, 27).Value = n
            C.Cells(i + j + 1, 27).Value = n
            End If

             If Abs(r.Cells(i, 1) - r.Cells(i + 2, 3)) < 0.003 And Abs(r.Cells(i, 2) - r.Cells(i + 2, 4)) <
0.003 And Abs(r.Cells(i, 3) - r.Cells(i + 1, 1)) < 0.003 And Abs(r.Cells(i, 4) - r.Cells(i + 1, 2)) < 0.003
Then
              C.Cells(i + 1, 27).Value = n
              C.Cells(i + 2, 27).Value = n
              C.Cells(i + 3, 27).Value = n
             End If
        j = j + 1
        n = n + 1
        Wend
        j = 1
      End If
      i = i + 1
      n = n + 1
   Wend

End Sub

Sub Trip()
   Dim q As Integer
   Dim i As Long
   Dim j As Long
```

```
Dim r As Range
Dim n As Long
Dim k As Long
Dim T As Worksheet
Dim C As Worksheet
Set T = ActiveWorkbook.Sheets("Cost")
Set C = ActiveWorkbook.Sheets("Trip")

' Dim Olat As String
' Dim Olong As String
' Dim DLat As String
' Dim DLong As String
q = 0
i = 1
j = 1
k = 2

While k < 30000
While C.Cells(k, 1) - C.Cells(k + 1, 1) = 0
Set r = Range("B2:E29080")
' While i < 300
    If C.Cells(i, 27).Value < 1 Then
        While j < 300
            If Abs(r.Cells(i, 1) - r.Cells(i + j, 1)) < 0.003 And Abs(r.Cells(i, 2) - r.Cells(i + j, 2)) < 0.003
And Abs(r.Cells(i, 3) - r.Cells(i + j, 3)) < 0.003 And Abs(r.Cells(i, 4) - r.Cells(i + j, 4)) < 0.003 Then
                If C.Cells(i + j + 1, 27).Value < 1 Then
                    C.Cells(i + j + 1, 27).Value = n
                    q = q + 1
```

```
            End If
            If C.Cells(i + 1, 27).Value < 1 Then
                C.Cells(i + 1, 27).Value = n
                q = q + 1
            End If
        End If

         If Abs(r.Cells(i, 1) - r.Cells(i + j, 3)) < 0.003 And Abs(r.Cells(i, 2) - r.Cells(i + j, 4)) < 0.003
And Abs(r.Cells(i, 3) - r.Cells(i + j, 1)) < 0.003 And Abs(r.Cells(i, 4) - r.Cells(i + j, 2)) < 0.003 Then
            If C.Cells(i + 1, 27).Value < 1 Then
                C.Cells(i + 1, 27).Value = n
            q = q + 1
            End If
            If C.Cells(i + j + 1, 27).Value < 1 Then
                 C.Cells(i + j + 1, 27).Value = n
            q = q + 1
            End If
         End If

          If Abs(r.Cells(i, 1) - r.Cells(i + 2, 3)) < 0.003 And Abs(r.Cells(i, 2) - r.Cells(i + 2, 4)) <
0.003 And Abs(r.Cells(i, 3) - r.Cells(i + 1, 1)) < 0.003 And Abs(r.Cells(i, 4) - r.Cells(i + 1, 2)) < 0.003
Then
             If C.Cells(i + 1, 27).Value < 1 Then
             C.Cells(i + 1, 27).Value = n
             q = q + 1
             End If
             If C.Cells(i + 2, 27).Value < 1 Then
             C.Cells(i + 2, 27).Value = n
```

```
                q = q + 1
                End If
                If C.Cells(i + 3, 27).Value < 1 Then
                C.Cells(i + 3, 27).Value = n
                q = q + 1
                End If
               ' C.Cells(i + 2, 27).Value = n
               ' C.Cells(i + 3, 27).Value = n
              End If
          j = j + 1
          'n = n + 1
          Wend
          j = 1
        End If
        C.Cells(i + 1, 28).Value = q
        i = i + 1
        k = k + 1
        n = n + 1
        q = 0
      'Wend


      'i = 1
      'j = 1
     Wend
      k = k + 1
   Wend


   End Sub
```

```
Sub NHTS()
```

Converting Lat/Long to distance:

```
Sub NHTS()

Dim i   As Long
```

```
Dim l1 As String

Dim l2 As String

Set A = ActiveWorkbook.Sheets("A")

i = 80000

l1 = 0.68059

l2 = 1.342265

While i < 200000

'If A.Cells(i - 1, 1).Value <> A.Cells(i, 1).Value Then

 ' l1 = A.Cells(i, 16).Value

 ' l2 = A.Cells(i, 17).Value

'End If

A.Cells(i, 26).Value = Application.Acos(Sin(l1) * Sin(A.Cells(i, 16).Value) + Cos(l1) * Cos(A.Cells(i,
16).Value) * Cos(A.Cells(i, 17).Value - l2)) * 3963.1676

'i = i + 1

If A.Cells(i + 1, 1).Value <> A.Cells(i, 1).Value Then

  l1 = A.Cells(i + 1, 16).Value

  l2 = A.Cells(i + 1, 17).Value

End If

i = i + 1

Wend

End Sub
```

Generating toll cost:

```
Sub Cost()


Dim i As Long
```

```
Dim j As Long
Dim e As Integer
Dim f As Integer
Dim h As Integer
Dim k As Integer
Dim m As Integer
Dim x As Integer
Dim A As Worksheet
Dim C As Worksheet
Set C = ActiveWorkbook.Sheets("Cost")
Set A = ActiveWorkbook.Sheets("After_full")
i = 2
j = 2
e = 0
f = 0
g = 0
h = 0
k = 0
x = 0
m = 0

While j < 872338



    If A.Cells(j, 18).Value = 1 Then
    d = 1


    End If
```

```
If A.Cells(j, 19).Value = 1 Then
e = 1
End If

If A.Cells(j, 20).Value = 1 Then
f = 1

End If

If A.Cells(j, 21).Value = 1 Then
g = 1
End If

If A.Cells(j, 22).Value = 1 Then
h = 1

End If

If A.Cells(j, 23).Value = 1 Then
k = 1
End If

'check weekends and holidays'
If A.Cells(j, 5).Value = 101211 Or A.Cells(j, 5).Value = 111211 Or A.Cells(j, 5).Value = 171211
Or A.Cells(j, 5).Value = 181211 Or A.Cells(j, 5).Value = 241211 Or A.Cells(j, 5).Value = 251211 Or
A.Cells(j, 5).Value = 261211 Or A.Cells(j, 5).Value = 311211 Or A.Cells(j, 5).Value = 10112 Or
A.Cells(j, 5).Value = 20112 Or A.Cells(j, 5).Value = 70112 Or A.Cells(j, 5).Value = 80112 Or
```

```
A.Cells(j, 5).Value = 140112 Or A.Cells(j, 5).Value = 150112 Or A.Cells(j, 5).Value = 210112 Or
A.Cells(j, 5).Value = 220112 Or A.Cells(j, 5).Value = 280112 Or A.Cells(j, 5).Value = 290112 Or
A.Cells(j, 5).Value = 40212 Or A.Cells(j, 5).Value = 50212 Or A.Cells(j, 5).Value = 110212 Or
A.Cells(j, 5).Value = 120212 Or A.Cells(j, 5).Value = 180212 Or A.Cells(j, 5).Value = 160112 Then
        m = 1
    End If

    If A.Cells(j, 10).Value > 0 Then



    '97 to 182'
        If f = 1 Then
         If A.Cells(j, 25) = 1 Then
         x = 1
          C.Cells(i, 12) = 0.6
          End If
          If A.Cells(j, 26) = 1 Or m = 1 Then
          x = 2
          C.Cells(i, 12) = 0.5
          End If
          If A.Cells(j, 27) = 1 Then
          x = 3
          C.Cells(i, 12) = 0.4
          End If
        End If

         '650 to 29'
```

```
If h = 1 Then
If A.Cells(j, 25) = 1 Then
x = 1
 C.Cells(i, 12) = 0.65
 End If
 If A.Cells(j, 26) = 1 Or m = 1 Then
 x = 2
 C.Cells(i, 12) = 0.55
 End If
 If A.Cells(j, 27) = 1 Then
 x = 3
 C.Cells(i, 12) = 0.4
 End If
End If


'29 to 95'
If k = 1 Then
If A.Cells(j, 25) = 1 Then
x = 1
 C.Cells(i, 12) = 0.7
 End If
 If A.Cells(j, 26) = 1 Or m = 1 Then
 x = 2
 C.Cells(i, 12) = 0.55
 End If
 If A.Cells(j, 27) = 1 Then
 x = 3
 C.Cells(i, 12) = 0.4
```

```
  End If
 End If

 '182 to 650'
If g = 1 Then
  If A.Cells(j, 25) = 1 Then
  x = 1
  C.Cells(i, 12) = 0.75
  End If
  If A.Cells(j, 26) = 1 Or m = 1 Then
  x = 2
  C.Cells(i, 12) = 0.6
  End If
  If A.Cells(j, 27) = 1 Then
  x = 3
  C.Cells(i, 12) = 0.4
  End If
 End If

  '97 to 650'
If f = 1 And g = 1 Then
  If A.Cells(j, 25) = 1 Then
  x = 1
  C.Cells(i, 12) = 1.3
  End If
  If A.Cells(j, 26) = 1 Or m = 1 Then
  x = 2
  C.Cells(i, 12) = 1.05
```

```
End If
If A.Cells(j, 27) = 1 Then
x = 3
C.Cells(i, 12) = 0.55
End If
End If

'182 to 29'
If g = 1 And h = 1 Then
If A.Cells(j, 25) = 1 Then
x = 1
C.Cells(i, 12) = 1.35
End If
If A.Cells(j, 26) = 1 Or m = 1 Then
x = 2
C.Cells(i, 12) = 1.1
End If
If A.Cells(j, 27) = 1 Then
x = 3
C.Cells(i, 12) = 0.55
End If
End If

'650 to 95'
If h = 1 And k = 1 Then
If A.Cells(j, 25) = 1 Then
x = 1
C.Cells(i, 12) = 1.35
```

```
      End If
      If A.Cells(j, 26) = 1 Or m = 1 Then
      x = 2
      C.Cells(i, 12) = 1.05
      End If
      If A.Cells(j, 27) = 1 Then
      x = 3
      C.Cells(i, 12) = 0.55
      End If
    End If

'370 to 97'
    If e = 1 Then
     If A.Cells(j, 25) = 1 Then
     x = 1
      C.Cells(i, 12) = 1.45
      End If
      If A.Cells(j, 26) = 1 Or m = 1 Then
      x = 2
      C.Cells(i, 12) = 1.15
      End If
      If A.Cells(j, 27) = 1 Then
      x = 3
      C.Cells(i, 12) = 0.6
      End If
    End If

     '97 to 29'
```

```
If f = 1 And h = 1 Then
    If A.Cells(j, 25) = 1 Then
    x = 1
     C.Cells(i, 12) = 1.95
     End If
     If A.Cells(j, 26) = 1 Or m = 1 Then
     x = 2
     C.Cells(i, 12) = 1.55
     End If
     If A.Cells(j, 27) = 1 Then
     x = 3
     C.Cells(i, 12) = 0.8
     End If
   End If


     '370 to 182'
   If e = 1 And f = 1 Then
    If A.Cells(j, 25) = 1 Then
    x = 1
     C.Cells(i, 12) = 2
     End If
     If A.Cells(j, 26) = 1 Or m = 1 Then
     x = 2
     C.Cells(i, 12) = 1.6
     End If
     If A.Cells(j, 27) = 1 Then
     x = 3
     C.Cells(i, 12) = 0.8
```

```
      End If
    End If
'182 to 95'
If g = 1 And k = 1 Then
     If A.Cells(j, 25) = 1 Then
     x = 1
      C.Cells(i, 12) = 2.05
      End If
      If A.Cells(j, 26) = 1 Or m = 1 Then
      x = 2
      C.Cells(i, 12) = 1.65
      End If
      If A.Cells(j, 27) = 1 Then
      x = 3
      C.Cells(i, 12) = 0.85
      End If
    End If

    '97 to 95'
If f = 1 And k = 1 Then
     If A.Cells(j, 25) = 1 Then
     x = 1
      C.Cells(i, 12) = 2.6
      End If
      If A.Cells(j, 26) = 1 Or m = 1 Then
      x = 2
      C.Cells(i, 12) = 2.1
      End If
```

```
  If A.Cells(j, 27) = 1 Then
  x = 3
  C.Cells(i, 12) = 1.05
  End If
End If


  '370 to 650'
If e = 1 And g = 1 Then
 If A.Cells(j, 25) = 1 Then
 x = 1
  C.Cells(i, 12) = 2.7
  End If
  If A.Cells(j, 26) = 1 Or m = 1 Then
  x = 2
  C.Cells(i, 12) = 2.2
  End If
  If A.Cells(j, 27) = 1 Then
  x = 3
  C.Cells(i, 12) = 1.1
  End If
End If
  '370 to 29'
If e = 1 And h = 1 Then
 If A.Cells(j, 25) = 1 Then
 x = 1
  C.Cells(i, 12) = 3.35
  End If
  If A.Cells(j, 26) = 1 Or m = 1 Then
```

```
    x = 2
    C.Cells(i, 12) = 2.7
    End If
    If A.Cells(j, 27) = 1 Then
    x = 3
    C.Cells(i, 12) = 1.35
    End If
   End If
    '370 to 97'
   If e = 1 And k = 1 Then
    If A.Cells(j, 25) = 1 Then
    x = 1
    C.Cells(i, 12) = 4
    End If
    If A.Cells(j, 26) = 1 Or m = 1 Then
    x = 2
    C.Cells(i, 12) = 3.2
    End If
    If A.Cells(j, 27) = 1 Then
    x = 3
    C.Cells(i, 12) = 1.6
    End If
   End If
C.Cells(i, 1).Value = A.Cells(j, 1).Value
'OUT.Cells(i, 2) = LD.Cells(j, 2)
'OUT.Cells(i, 3) = LD.Cells(j, 3)
C.Cells(i, 2).Value = A.Cells(j, 2).Value
C.Cells(i, 3).Value = A.Cells(j, 15).Value
```

```
C.Cells(i, 21).Value = A.Cells(j - (A.Cells(j, 14).Value), 2).Value
C.Cells(i, 22).Value = A.Cells(j - (A.Cells(j, 14).Value), 15).Value
'OUT.Cells(i, 6) = LD.Cells(j, 6)
'OUT.Cells(i, 7) = LD.Cells(j, 7)
'OUT.Cells(i, 8) = LD.Cells(j, 8)
'OUT.Cells(i, 9) = LD.Cells(j, 9)
C.Cells(i, 4).Value = A.Cells(j, 4).Value
'OUT.Cells(i, 11) = LD.Cells(j, 11)
C.Cells(i, 5).Value = A.Cells(j, 16).Value
C.Cells(i, 6).Value = A.Cells(j, 17).Value
C.Cells(i, 7).Value = A.Cells(j, 32).Value
C.Cells(i, 8).Value = A.Cells(j, 14).Value
C.Cells(i, 9).Value = x
C.Cells(i, 10).Value = m
C.Cells(i, 13).Value = e
C.Cells(i, 14).Value = f
C.Cells(i, 15).Value = g
C.Cells(i, 16).Value = h
C.Cells(i, 17).Value = k
C.Cells(i, 18).Value = A.Cells(j, 25).Value
C.Cells(i, 19).Value = A.Cells(j, 26).Value
C.Cells(i, 20).Value = A.Cells(j, 27).Value
i = i + 1
x = 0
e = 0
f = 0
g = 0
h = 0
```

```
    k = 0
    End If
  j = j + 1
  m = 0
  Wend
End Sub
```

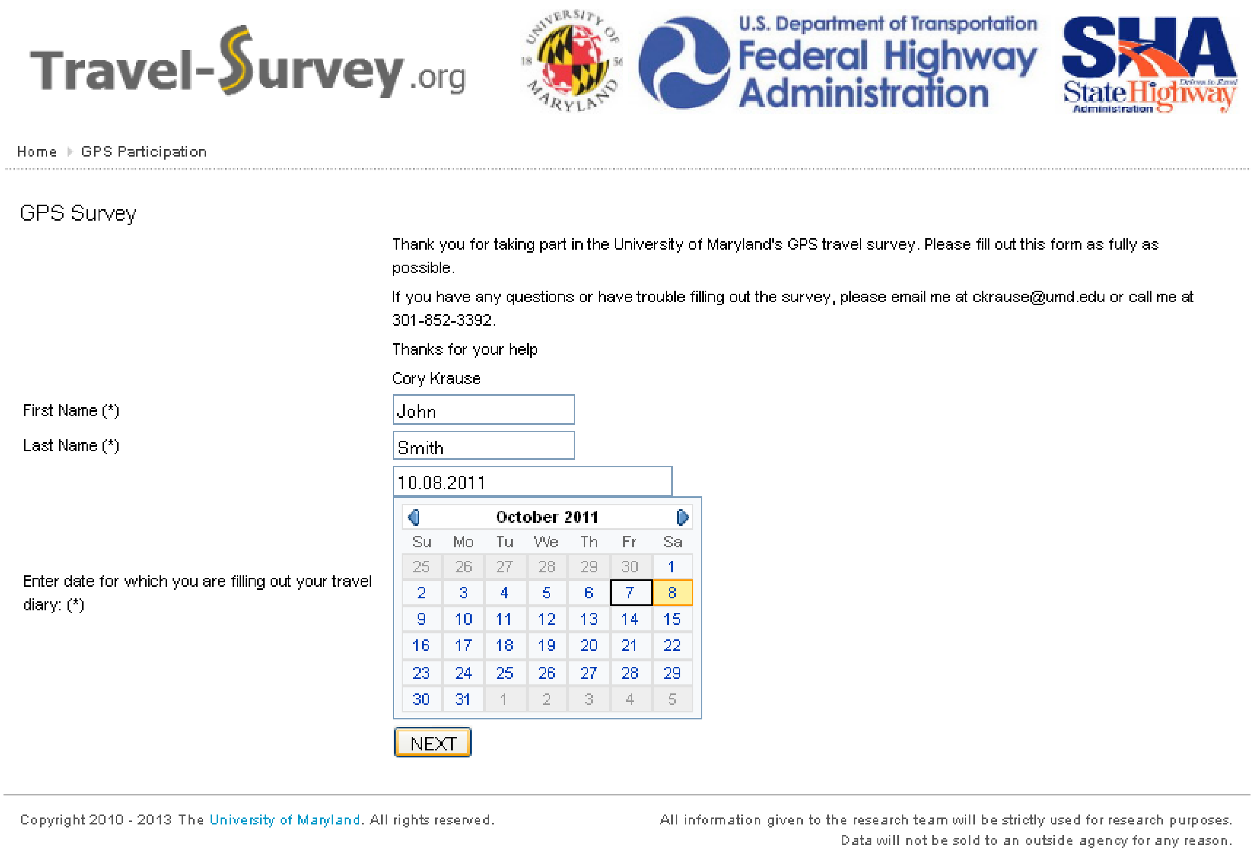
Appendix B: GPS Forms and information mailed to survey participants:

Travel Diary

Travel Diary Example:

This form is a simple example of how to fill out the travel survey. If you are having trouble understanding the layout or how to fill out the diary, please contact Cory Krause at ckrause@umd.edu with your questions.

First I input my name and for which day I am filling out the survey for.



Travel-Survey.org

U.S. Department of Transportation Federal Highway Administration

SHA State Highway Administration

Home GPS Participation

GPS Survey

Thank you for taking part in the University of Maryland's GPS travel survey. Please fill out this form as fully as possible.

If you have any questions or have trouble filling out the survey, please email me at ckrause@umd.edu or call me at 301-852-3392.

Thanks for your help

Cory Krause

First Name (*) John

Last Name (*) Smith

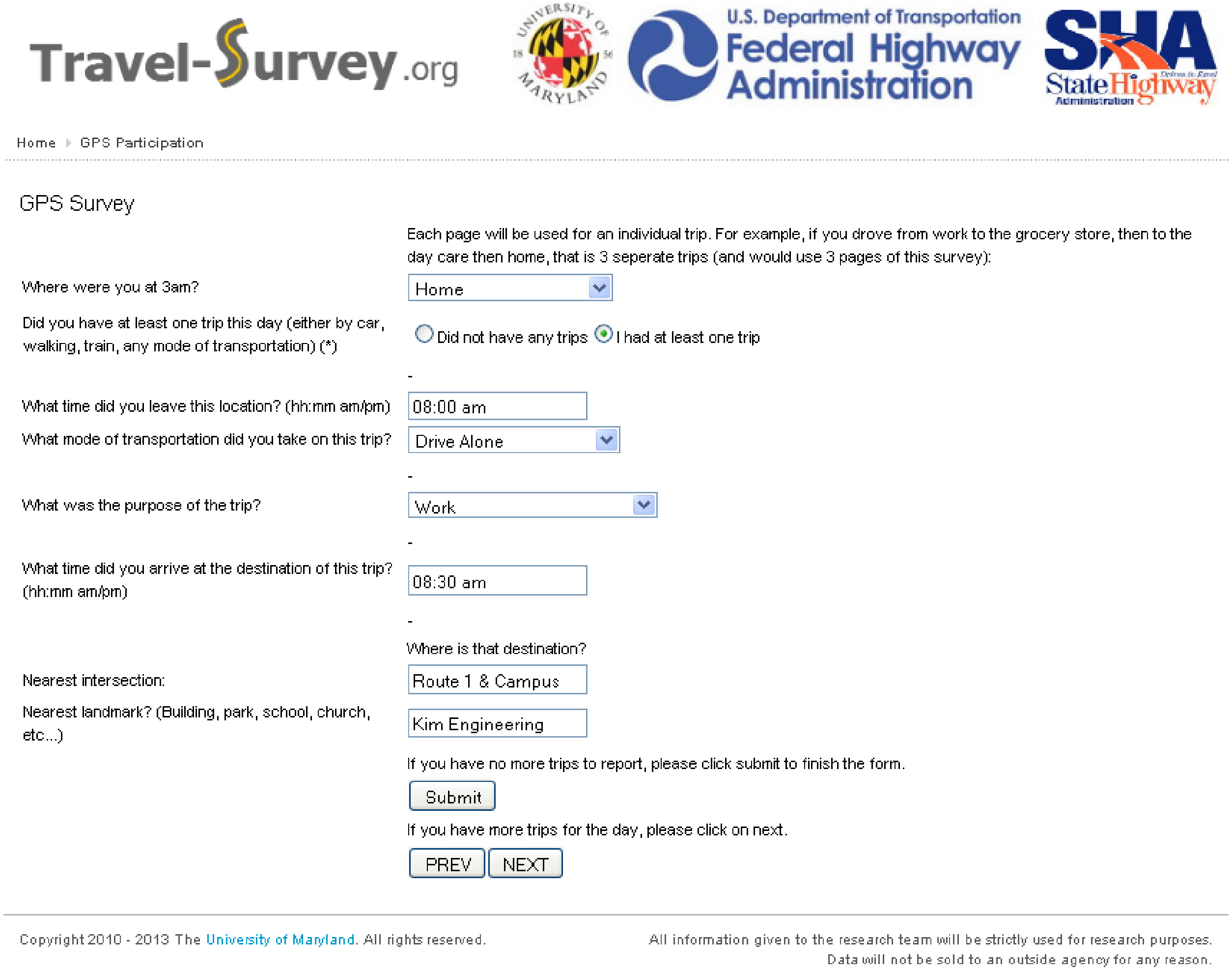
Enter date for which you are filling out your travel diary: (*) 10.08.2011

October 2011

Su	Mo	Tu	We	Th	Fr	Sa
25	26	27	28	29	30	1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	1	2	3	4	5

NEXT

All information given to the research team will be strictly used for research purposes. Data will not be sold to an outside agency for any reason.


Travel-Survey.org
U.S. Department of Transportation
Federal Highway Administration
SHA
State Highway Administration

Home ▸ GPS Participation

GPS Survey

Each page will be used for an individual trip. For example, if you drove from work to the grocery store, then to the day care then home, that is 3 seperate trips (and would use 3 pages of this survey):

Where were you at 3am? Home

Did you have at least one trip this day (either by car, walking, train, any mode of transportation) (*) ○ Did not have any trips ◉ I had at least one trip

What time did you leave this location? (hh:mm am/pm) 08:00 am

What mode of transportation did you take on this trip? Drive Alone

What was the purpose of the trip? Work

What time did you arrive at the destination of this trip? (hh:mm am/pm) 08:30 am

Where is that destination?

Nearest intersection: Route 1 & Campus

Nearest landmark? (Building, park, school, church, etc...) Kim Engineering

If you have no more trips to report, please click submit to finish the form.

Submit

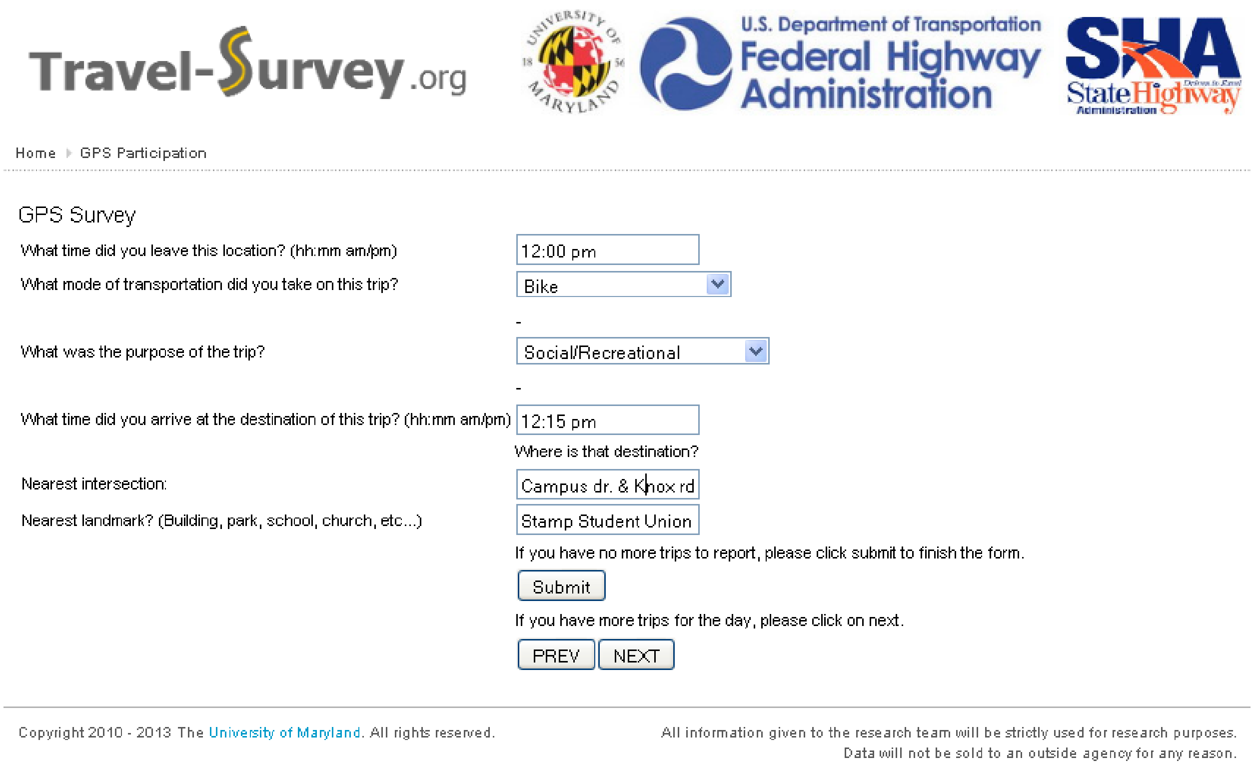
If you have more trips for the day, please click on next.

PREV NEXT

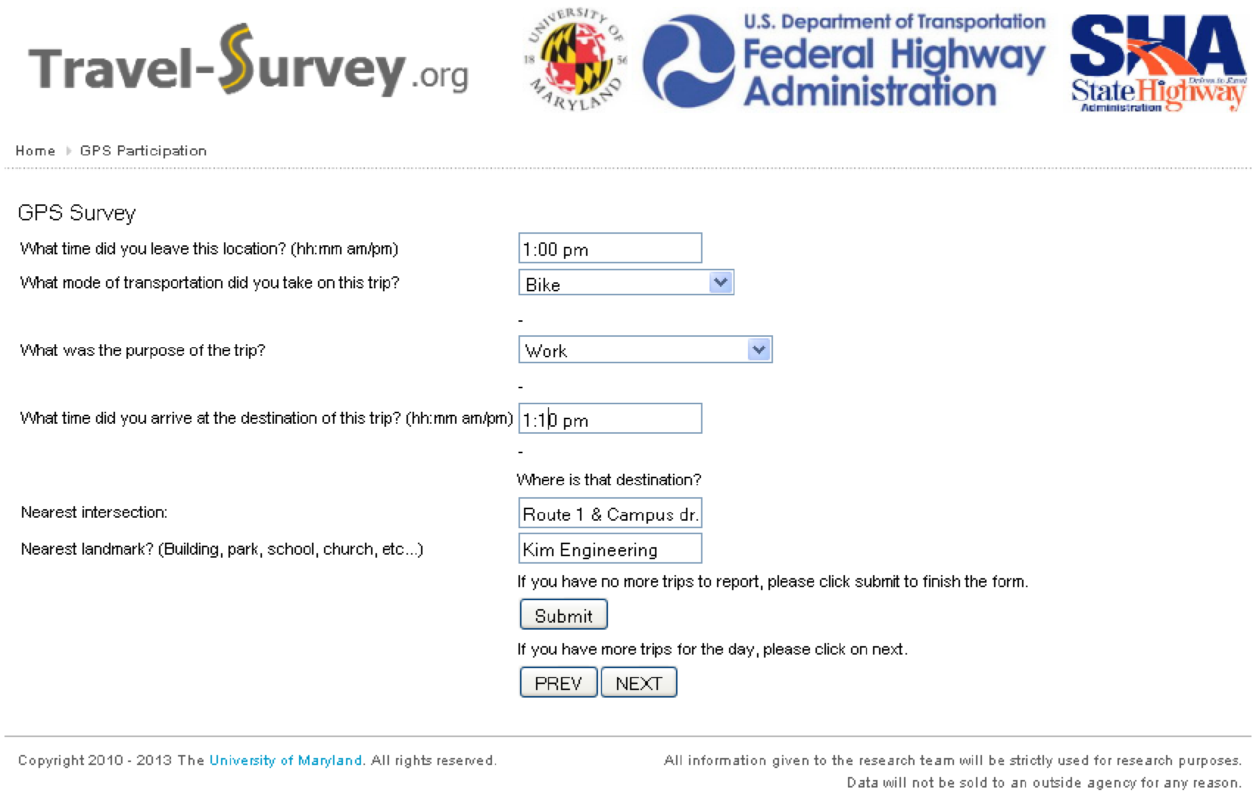
Copyright 2010 - 2013 The University of Maryland. All rights reserved.

All information given to the research team will be strictly used for research purposes. Data will not be sold to an outside agency for any reason.

On the following page I fill out my first trip of the day: my drive to work. I leave at 8 am and arrive at 8:30 am. I fill out the nearest intersection as well as the nearest landmark. Since I have another trip for the day, I click ‘NEXT’.



Here I fill out my trip to get lunch. I leave at 12 pm (noon) and arrive at my location at 12:15pm. I rode my bike there and it is classified as social/recreational.



Travel-Survey.org

U.S. Department of Transportation Federal Highway Administration

SHA State Highway Administration

Home › GPS Participation

GPS Survey

What time did you leave this location? (hh:mm am/pm) 1:00 pm

What mode of transportation did you take on this trip? Bike

What was the purpose of the trip? Work

What time did you arrive at the destination of this trip? (hh:mm am/pm) 1:10 pm

Where is that destination?

Nearest intersection: Route 1 & Campus dr.

Nearest landmark? (Building, park, school, church, etc...) Kim Engineering

If you have no more trips to report, please click submit to finish the form.

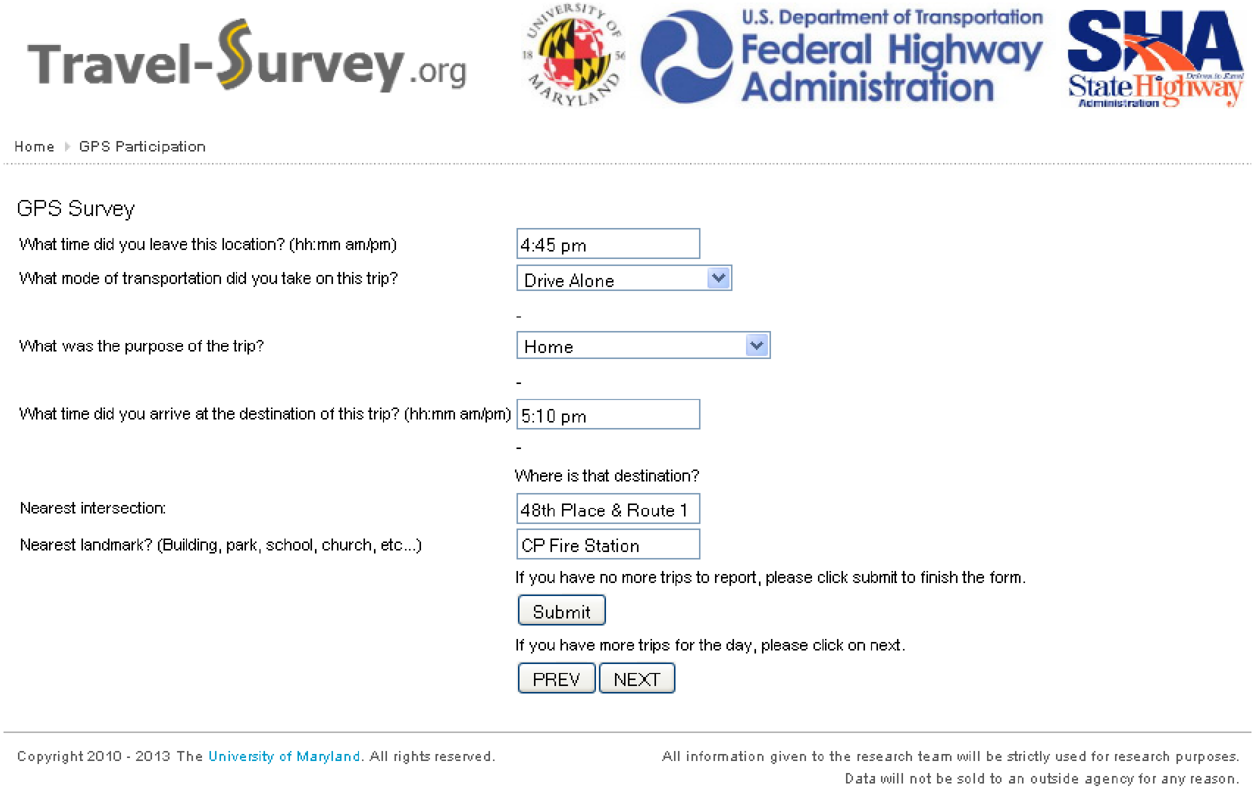
Submit

If you have more trips for the day, please click on next.

PREV NEXT

All information given to the research team will be strictly used for research purposes. Data will not be sold to an outside agency for any reason.

Shown above is my return trip from lunch.



Home ▸ GPS Participation

GPS Survey

What time did you leave this location? (hh:mm am/pm) 4:45 pm

What mode of transportation did you take on this trip? Drive Alone

What was the purpose of the trip? Home

What time did you arrive at the destination of this trip? (hh:mm am/pm) 5:10 pm

Where is that destination?

Nearest intersection: 48th Place & Route 1

Nearest landmark? (Building, park, school, church, etc...) CP Fire Station

If you have no more trips to report, please click submit to finish the form.

Submit

If you have more trips for the day, please click on next.

PREV NEXT

Copyright 2010 - 2013 The University of Maryland. All rights reserved.

All information given to the research team will be strictly used for research purposes. Data will not be sold to an outside agency for any reason.

And finally, my trip home for the day. If this is your last trip of the day, please click Submit to save the form.

Again, if you have any questions, feel free to call or email me. Thank you for your help.

Cory Krause

Transportation Systems Research Lab

University of Maryland

ckrause@umd.edu

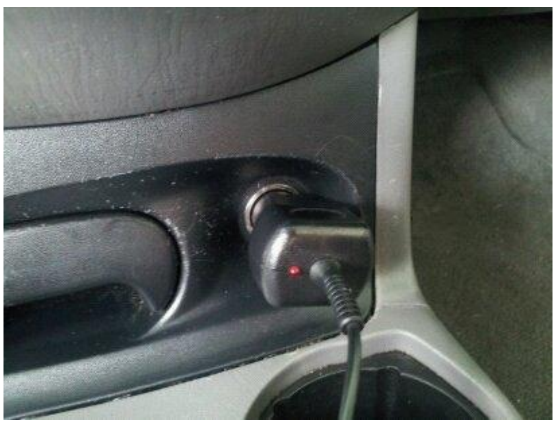
301-852-3392

GPS Installation

GPS Installation Instructions

Thank you for taking our online travel survey and taking part in the GPS portion of our project. This document will show you how to install the GPS device in your vehicle. If at any time you have a problem with the directions, don't hesitate to email me at ckrause@umd.edu

Step 1: Plug the car charger into the vehicle that you take on a regular basis (most used vehicle) via the cigarette lighter receptacle.



Step 2: Insert the mini USB end of the car charger in the GPS port on the side of the device.

Step3: Move the switch on the side of the GPS all the way to the left so that it is on the "LOG" section.

Step 4: Drive normally and do not touch any of the buttons on the device during your travel survey period. When not driving, leave the device in the car, it will go into sleep mode after not moving for a few minutes. After the trial period is over, simply turn off the device, unplug it, and return using the shipping label we will have given you.

Step 5: Twice during the trial period, I will send you an email that asks you to validate your data. Please see the included document for an example of how to fill out the travel survey.

Thanks again for your help,

Cory Krause

Transportation Systems Research

University of Maryland

ckrause@umd.edu

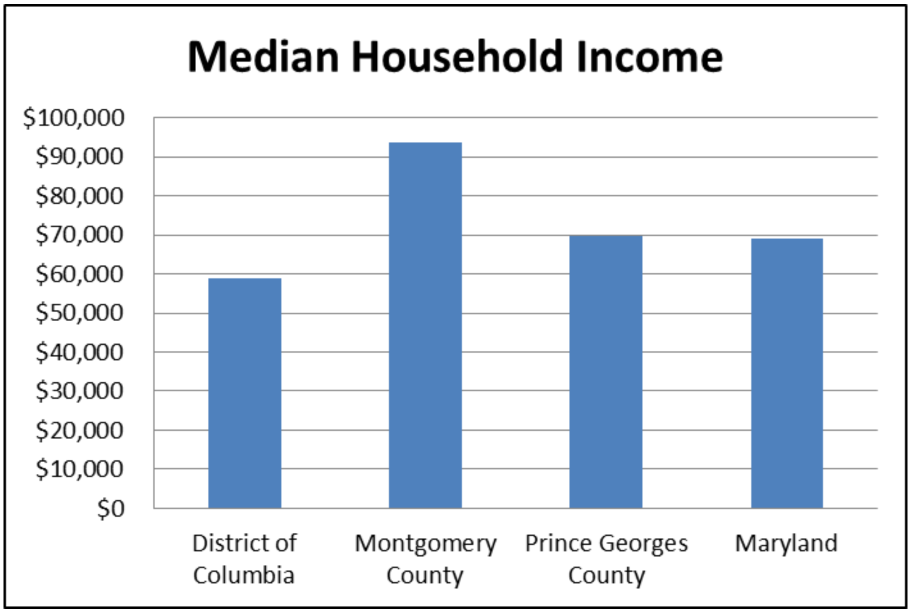
301-852-3392

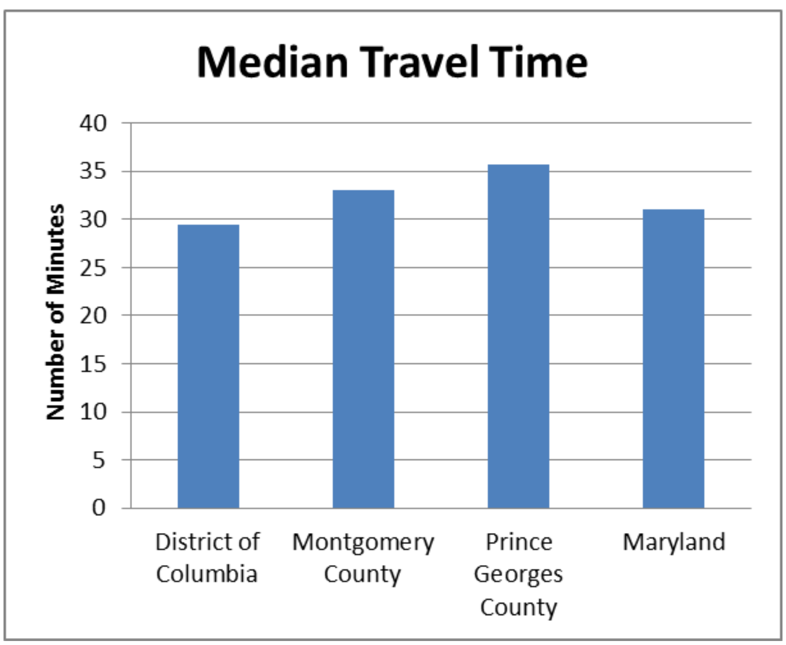
Return Shipment Instructions:

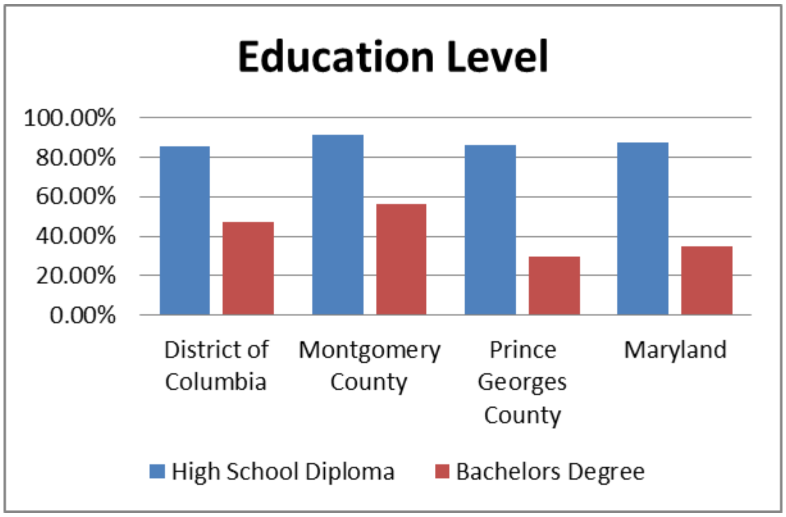
- Fill out W-9 form
- Put GPS device and charger in the box. (make sure the GPS device is turned off)
- Sign Honorarium stating you participated in a Survey. Fill out all information including mailing address.
- Sign and date the consent form
- Take the return shipment label from inside the box and adhere it to the outside of the box. You only need to fill out section 1 of the form (your address information). All other sections can be left blank or are already filled out.
- Make sure that previous shipping labels are removed or covered by the new shipping label.
- Ship the package from any FedEx location. (All shipping costs have been prepaid).

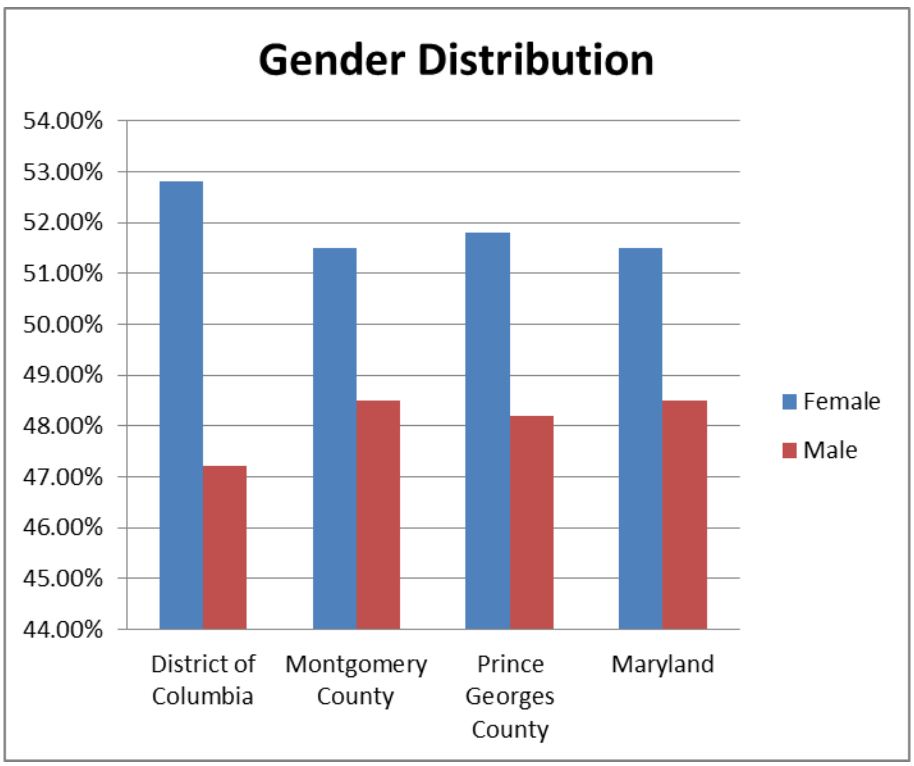
Appendix C, GPS Survey demographic information:

Surrounding Area

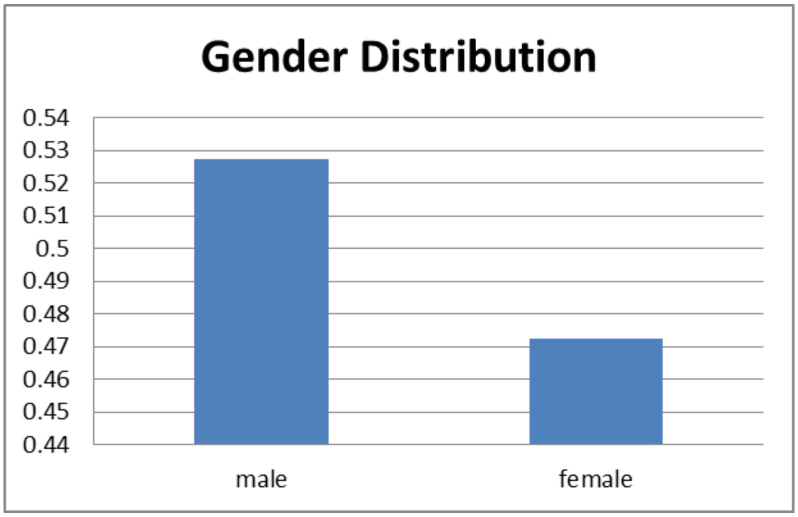


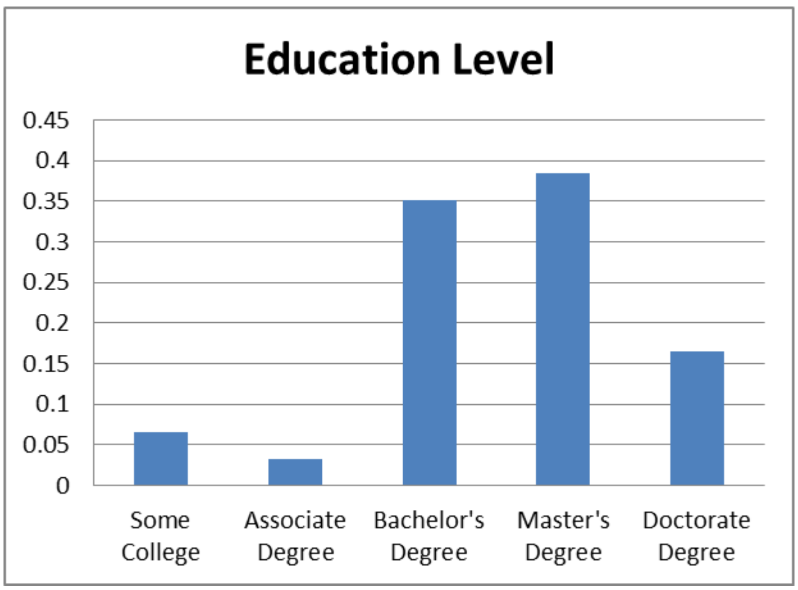


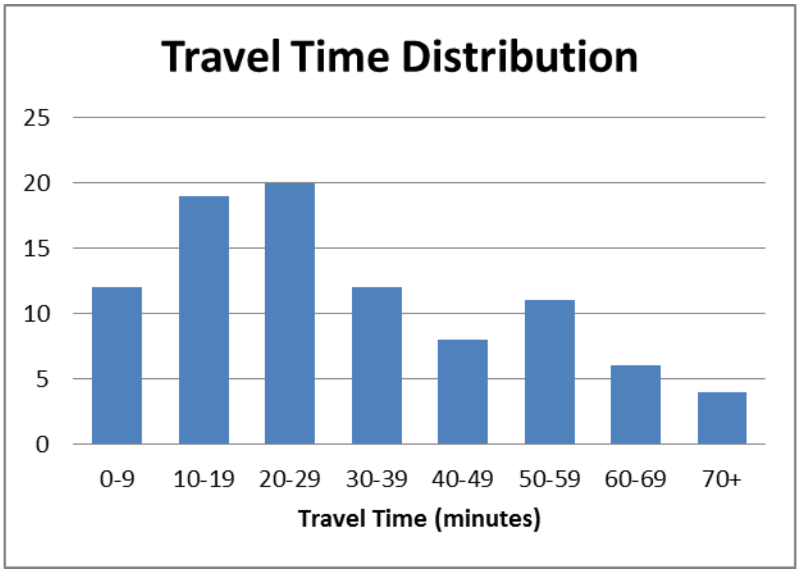

Education Level
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
District of Columbia
Montgomery County
Prince Georges County
Maryland
High School Diploma
Bachelors Degree

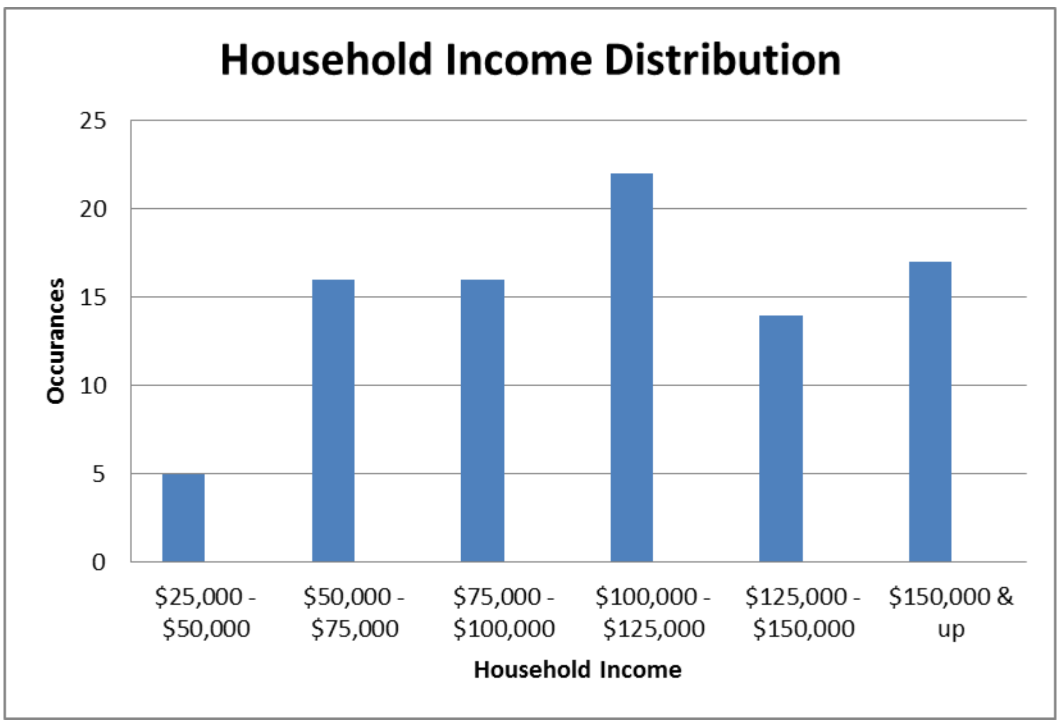

Gender Distribution
54.00%
53.00%
52.00%
51.00%
50.00%
49.00%
48.00%
47.00%
46.00%
45.00%
44.00%
District of Columbia
Montgomery County
Prince Georges County
Maryland
Female
Male

Pilot Study Participants






Travel Time Distribution
25
20
15
10
5
0
0-9
10-19
20-29
30-39
40-49
50-59
60-69
70+
Travel Time (minutes)


Household Income Distribution
Occurances
25
20
15
10
5
0
$25,000 - $50,000
$50,000 - $75,000
$75,000 - $100,000
$100,000 - $125,000
$125,000 - $150,000
$150,000 & up
Household Income

Appendix D: Positive Model Raw Outputs

Search rules

=== Run information ===

Scheme:weka.classifiers.rules.JRip -F 3 -N 2.0 -O 2 -S 1

Relation: decision_switch_rules-weka.filters.unsupervised.attribute.Remove-R1,8-

weka.filters.unsupervised.attribute.Remove-R7-

weka.filters.unsupervised.attribute.Remove-R7,13,15-16-

weka.filters.unsupervised.attribute.Remove-R28-

weka.filters.unsupervised.attribute.Remove-R26-

weka.filters.unsupervised.attribute.Remove-R24-25-

weka.filters.unsupervised.attribute.Remove-R27-

weka.filters.unsupervised.attribute.Remove-R25-

weka.filters.unsupervised.attribute.Remove-R22-23-

weka.filters.unsupervised.attribute.Remove-R20

Instances: 924

Attributes: 27

Cost_Holiday

Peak

Off-Peak

Night

Age

Sex

d_time

First_OD

d_distance

d_cost

d_VOT

average_time_combined

value_of_time_true

distance_on_icc

time_if_ICC

Distance

Cost

ICC_trip_distance_percent

yesterday_trip

yesterday_speed

Income

no_icc_time

No_icc_time

Icc_time

ICC_assumed_savings

Search

Segments

Test mode:10-fold cross-validation

=== Classifier model (full training set) ===

JRIP rules:

===========

1(d_VOT >= -16.5) and (distance_on_icc <= 7.72) and (value_of_time_true <= 5.5) and (yesterday_trip = 0) => Search=0 (112.0/16.0)

2(d_VOT <= 4.6) and (Age >= 58) and (d_VOT >= -6.128571) => Search=0 (35.0/6.0)

3(ICC_trip_distance_percent <= 0.052018) and (no_icc_time <= 64.110921) => Search=0 (24.0/0.0)

4(Sex = Female) and (d_VOT >= -1.185829) and (distance_on_icc >= 10.1) and (d_VOT <= 1.536147) => Search=0 (13.0/1.0)

5(average_time_combined <= 22.444444) and (Age <= 31) and (average_time_combined >= 20.285714) => Search=0 (12.0/1.0)

=> Search=1 (728.0/50.0)

Number of Rules : 6

Time taken to build model: 0.23 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	816	88.3117 %
Incorrectly Classified Instances	108	11.6883 %
Kappa statistic	0.6697	
Mean absolute error	0.161	
Root mean squared error	0.3204	
Relative absolute error	44.0659 %	
Root relative squared error	74.9814 %	
Total Number of Instances	924	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.712	0.063	0.782	0.712	0.745	0.83	0
	0.937	0.288	0.911	0.937	0.924	0.83	1
Weighted Avg.	0.883	0.234	0.88	0.883	0.881	0.83	

=== Confusion Matrix ===

```
  a   b   <-- classified as
 158  64 |   a = 0
  44 658 |   b = 1
```

Route Decision rules:

=== Run information ===

Scheme:weka.classifiers.rules.JRip -F 3 -N 2.0 -O 2 -S 1

Relation: decision_switch_rules-weka.filters.unsupervised.attribute.Remove-R1,8,10,16-19,22,29-30,35,39-40-weka.filters.unsupervised.attribute.Remove-R21-22-weka.filters.unsupervised.attribute.Remove-R21-weka.filters.unsupervised.attribute.Remove-R23-24-weka.filters.unsupervised.attribute.Remove-R22-weka.filters.unsupervised.attribute.Remove-R19-weka.filters.unsupervised.attribute.Remove-R13-weka.filters.unsupervised.attribute.Remove-R10-weka.filters.unsupervised.attribute.Remove-R10-weka.filters.unsupervised.attribute.Remove-R8-10

Instances: 924

Attributes: 15

Cost_Holiday

Peak

Off-Peak

Night

Age

Sex

ICC

distance_on_icc

Distance

Cost

ICC_trip_distance_percent

yesterday_trip

yesterday_speed

Income

ICC_assumed_savings

Test mode:10-fold cross-validation

=== Classifier model (full training set) ===

JRIP rules:

===========

(Age <= 37) and (ICC_assumed_savings <= -11.978792) and (ICC_trip_distance_percent <= 0.859017) => ICC=1 (49.0/3.0)

(distance_on_icc >= 7.86) and (Age <= 42) and (ICC_trip_distance_percent <= 0.67612) and (ICC_trip_distance_percent >= 0.536169) => ICC=1 (18.0/0.0)

(Distance >= 31.141895) and (ICC_trip_distance_percent >= 0.246393) => ICC=1 (30.0/6.0)

(yesterday_speed <= 0.188022) and (ICC_assumed_savings >= 14.079605) => ICC=1 (7.0/0.0)

(ICC_assumed_savings <= -31.226118) and (Distance >= 8.677326) and (Cost >= 0.8) => ICC=1 (15.0/1.0)

=> ICC=0 (805.0/144.0)

Number of Rules : 6

Time taken to build model: 0.12 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	741	80.1948 %
Incorrectly Classified Instances	183	19.8052 %
Kappa statistic	0.436	
Mean absolute error	0.2921	
Root mean squared error	0.4006	
Relative absolute error	73.4109 %	
Root relative squared error	89.8388 %	
Total Number of Instances	924	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.934	0.549	0.819	0.934	0.873	0.687	0
	0.451	0.066	0.722	0.451	0.555	0.687	1
Weighted Avg.	0.802	0.417	0.792	0.802	0.786	0.687	

=== Confusion Matrix ===

```
  a   b   <-- classified as
 627  44 |   a = 0
 139 114 |   b = 1
```

Bibliography

1. Anderson, R., and Costinett, P. (2007). “Ohio Statewide Passenger Transit Calibration”, presented at the 11th TRB National Transportation Planning Applications Conference, Daytona Beach, Florida.
2. Arentze, T. A. and Timmermans, H.J.P. (2000). ALBATROSS: A Learning-Based Transportation Oriented Simulation System, Eindhoven, EIRSS.
3. Arentze, T. A. and Timmermans, H.J.P. (2003). Modeling learning and adaptation processes in activity-travel choice: A framework and numerical experiment, Transportation 30: 37–62.
4. Arentze, T. A. and Timmermans, H.J.P. (2004). A learning-based transportation oriented simulation system, Transportation Research Part B 38: 613–633.
5. Arentze, T. A. and Timmermans, H.J.P. (2005). Information gain, novelty seeking and travel: a model of dynamic activity-travel behavior under conditions of uncertainty, Transportation Research Part A 39: 125–145.
6. Arentze, T. A. and Timmermans, H.J.P. (2005). Modeling learning and adaptation in transportation contexts, Transportmetrica 1(1): 13–22.
7. Arentze, T., Pelizaro, C. and Timmermans, H. (2005). Implementation of a model of dynamic activity-travel rescheduling decisions: an agent-based micro-simulation framework, Proceedings of CUPUM 05,

Computers in Urban Planning and Urban Management, 30-Jun-2005, London, CASA Centre for Advanced Spatial Analysis -University College London, London, pp. paper– 48.

8. Ashiabor, S., Baik, H., & Trani, A. (2007). Logit models to forecast nationwide intercity travel demand in the United States. Journal of the Transportation Research Record, 2007, 1–12.

9. Atlanta Vehicle Instrumented GPS data: Destination Choice Behavior and Activity Spaces, *ArbeitsberichteVerkehrs- und Raumplanung*, **303**, IVT, ETH, Zürich.

10. Auld, J., Mohammadian K., Doherty, S. (2008). Analysis of activity conflict resolution strategies, 1-24.

11. Auld, J., Williams, C., Mohammadian, K., Nelson, P. (2008). AN AUTOMATED GPS-BASED PROMPTED RECALL SURVEY WITH LEARNING ALGORITHMS, *500*.

12. Becker, G. S. (1965) A Theory of the Allocation of Time.pdf.

13. Ben-Akiva, M. and de Palma, A. (1986). Analysis of a dynamic residential location choice model with transaction costs. Journal of Regional Science, 26: 321–341.

14. Ben-Akiva, M., de Palma, A. and Kaysi, I. (1991). Dynamic network models and driver information systems, Transportation Research Part A 25(5): 251–266.

15. Bricka, S. (n.d.). Variations in Long-Distance Travel, 197-206. *TRB Transportation Research Circular E-C026—Personal Travel: The Long and Short of It*

16. Bureau of Transportation Statistics. (1997). 1995 American travel survey. Washington, DC: US Department of Transportation. Retrieved February 11, 2009, from http://www.bts.gov/publications/1995_american_travel_survey/

17. Burris, M. W., & Pendyala, R. M. (2002). Discrete choice models of traveler participation in differential time of day pricing programs. *Transport Policy*, *9*(3), 241-251. doi:10.1016/S0967-070X(02)00002-1

18. Burris, M. W., & Stockton, B. R. (n.d.). HOT Lanes in Houston — Six Years of Experience, 1-21.

19. Burris, M. W., Konduru, K. K., & Swenson, C. R. (2004). Long-Run Changes in Driver Behavior Due to Variable Tolls, (1864), 78-85.

20. Cambridge Systematics. (2008). National travel demand forecasting model phase I final scope. NCHRP Project 08-36, Task 70, National Cooperative Highway Research Program. Washington, DC: Transportation Research Board.

21. Cao, L., & Krumm, J. (2009). From GPS traces to a routable road map. *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems - GIS ’09*, 3. New York, New York, USA: ACM Press. doi:10.1145/1653771.1653776

22. Chen, B. (2011). Study of Interregional Long-Distance Commuting Using NHTS data.
23. Cirillo, C., & Axhausen, K. W. (2006). Evidence on the distribution of values of travel time savings from a six-week diary. *Transportation Research Part A: Policy and Practice*, *40*(5), 444–457. doi:10.1016/j.tra.2005.06.007
24. Costinett, P. and Stryker, A. (2007). “Calibrating the Ohio Statewide Travel Model”,presented at the 11th TRB National Transportation Planning Applications Conference, Daytona Beach, Florida.
25. Emmerink, R. H. M., Axhausen, K. W., Njikamp, P. and Rietveld, P. (1995). The potentil of information provision in a simulated road transport network with non-recurrent congestion, Transportation Research Part C 3(5): 293–309.
26. Erhardt, G. D., Freedman, J., Stryker, A., Fujioka, H., & Anderson, R. (2007). Ohio Long-Distance Travel Model, (2003), 130-138. doi:10.3141/2003-16
27. Erhardt, G., & Freedman, J. (n.d.). CALIBRATION AND APPLICATION OF THE OHIO LONG DISTANCE TRAVEL MODEL.
28. Erhardt, G., Freedman, J., Stryker, A., Fujioka, H. and Anderson, R. (2007). “The Ohio Long Distance Travel Model”, accepted for publication in *Transportation Research Record*, Washington, D.C.
29. Giaimo, G., R. Anderson, L. Wargelin, and P. Stopher. Will It Work? Pilot Results from First Large-Scale Global Positioning System–Based

Household Travel Survey in the United States. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2176,* Transportation Research Board of the National Academies, Washington, D.C.,

30. Giaimo, G.T.,& Schiffer, R. (Eds.). (2005, August). Statewide travel demand modeling: A peer exchange. Transportation Research Circular, #E-C075. Washington, DC: Transportation Research Board.
31. Hariharan, R., Krumm, J., & Horvitz, E. (2005). Web-Enhanced GPS, T. Strang and C. Linnhoff-Popien (Eds.): LoCA 2005, LNCS 3479, pp. 95 – 104, 2005.
32. Hess, S., Bierlaire, M., & Polak, J. W. (2005). Estimation of value of travel-time savings using mixed logit models. *Transportation Research Part A: Policy and Practice*, *39*(2-3), 221-236. doi:10.1016/j.tra.2004.09.007
33. Horowitz, J. L. (1984). The stability of stochastic equilibrium in a two-link transportation network, Transportation Research Part B 18(1): 13–28.
34. Lapparent, M. D. (2002). Nonlinearities in the Valuations of Travel Times at the Individual Level .
35. Lei Zhang, Frank Southworth, Chenfeng Xiong & Anthon Sonnenberg (2012): Methodological Options and Data Sources for the Development of Long-Distance Passenger Travel Demand Models: A

Comprehensive Review, Transport Reviews: A Transnational Transdisciplinary Journal, 32:4, 399-433

36. Li, H., Guensler, R., & Ogle, J. (2005). An Analysis of Morning Commute Route Choice Patterns Using GPS Based Vehicle Activity Data. *TRB*, *30332*(404).

37. Mallett, W. (n.d.). Long-Distance Travel by Low-Income Households. *TRB Transportation Research Circular E-C026—Personal Travel: The Long and Short of It (169-177)*

38. McGucken, N. (2006). *NPTS brief*. http://nhts.ornl.gov/briefs/Long%20Distance%20Travel.pdf

39. Miller, E.J. (2004). The trouble with intercity travel demand models. Transportation Research Record, 1895, 94–101.

40. Murakami, E. "Can using global positioning system (GPS) improve trip reporting?." Transportation research. Part E, Logistics and transportation review 7.2-3 (1999):149.

41. Murakami, E., & Young, J. (1997). Daily Travel by Persons with Low Income Daily Travel by Persons with Low Income.

42. Nakayama, S., Kitamura, R. and Fujii, S. (2001). Drivers' route choice rules and network behavior: Do drivers become rational and homogeneous through learning?, Transportation Research Record: Journal of the Transportation Research Board 1752: 62–68.

43. Oliveira, M. G. S., Vovsha, P., Wolf, J., Birotker, Y., Givon, D., & Paasche, J. (2011). Global Positioning System-Assisted Prompted

Recall Household Travel Survey to Support Development of Advanced Travel Model in Jerusalem, Israel. *Transportation Research Record: Journal of the Transportation Research Board*, *2246*(-1), 16-23. doi:10.3141/2246-03\

44. Outwater, M., Sall, E., Tierney, K., Bradley, M., Kuppam, A., & Modugula, V. (n.d.). California Statewide Model for High-Speed Rail, *3*(1), 58-83.

45. Picado, R., Freedman, J., Stryker, A. and Erhardt, G. (2007). "Design, Estimation and Calibration of the Ohio Statewide Short Distance Travel Models", presented at the 11th TRB National Transportation Planning Applications Conference, Daytona Beach, Florida.

46. Prato, C. G. (2009). Route choice modeling : past , present and future research directions, *2*(1), 65-100.

47. Richardson, A., SEETHALER, R. (n.d.). Estimating Long-Distance Travel Behavior from the Most Recent Trip. *TRB Transportation Research Circular E-C026—Personal Travel: The Long and Short of It*

48. Satellite, T., & Programme, A. (2011). International Recommendations for Tourism Statistics 2008 Draft Compilation Guide, (March).

49. Schönfelder, S., H. Li, R. Guensler, J. Ogle and K.W. Axhausen (2005) Analysis of Commute

50. Small, K., Winston, C., & Yan, J. (2005). Uncovering the distribution of motorists' preferences for travel time and reliability b, *73*(4), 1367-1382.

51. Southworth, F., & Hu, P.S. (2010). The American long distance personal travel data and modeling program: A roadmap. Revised draft. Prepared for the Federal Highway Administration, Washington, DC.
52. Wolf, J. (2000). Using GPS Data Loggers To Replace Travel Diaries In the Collection of Travel Data by, (July).
53. Wolf, J., Guensler, R., Bachman, W. Elimination of the Travel Diary : An Experiment to Derive Trip Purpose From GPS Travel Data, (01).
54. Wolf, J., Loechl, M., Myers, J., Arce, C. TRIP RATE ANALYSIS IN GPS-ENHANCED PERSONAL TRAVEL By, (August 2001).
55. Wolf, J., M. Loechl, M. Thompson, and C. Arce. Trip Rate Analysis in GPS-Enhanced Personal Travel Surveys. In *Transport Survey Quality and Innovation* (P. Stopher and P. M. Jones, eds.), Pergamon, Oxford, United Kingdom, 2003, pp. 483–498.
56. World Tourism Organization. (2011). International Recommendations for Tourism Statistics 2008 Draft Compilation Guide, (March).
57. Xu, Y., Zuyeva, L., Kall, D., Elango, V., & Guensler, R. (2009). Mileage Based Value Pricing: Phase II Case Study Implications of Commute Atlanta Project. *TRB*, (18).
58. Zhang, L. (2006a). Search, information, learning and knowledge in travel-decision making. Ph.D. Dissertation, Department of Civil Engineering, University of Minnesota.

59. Zhang, L. (2006b). An agent-based behavioral model of spatial learning and route choice. In the compendium of papers CD of the 85th Transportation Research Board Annual Meetings, Washington, DC.

60. Zhang, L. (2006c). Traffic diversion effect of ramp metering at individual and system levels. *Transportation Research Record: Journal of the Transportation Research Board*, 2012: 20-29.

61. Zhang, L. (2007). Developing a positive approach to travel demand analysis: Silk theory and behavioral user equilibrium, in R. E. Allsop, B. M. G. H. and B. G. Heydecker (eds), Transportation and Traffic Theory, Elsevier, chapter 34, pp. 791–812.

62. Zhang, L. and Levinson, D. (2004). Agent-Based Approach to Travel Demand Modeling. In *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1898, 2004, pp. 28–36.

63. Zhang, L., Levinson, D., and Zhu, S. (2008). Agent-Based Model of Price Competition, Capacity Choice, and Product Differentiation on Congested Networks (August 1, 2008). *Journal of Transport Economics and Policy*, 42(3): 435-461

64. Zhang, L., Xiong, C., & Berger, K. (2010). Multimodal inter-regional origin-destination demand estimation: A review of methodologies and their applicability to national-level travel analysis in the US WCTR. Lisbon, Portugal, July, 2010.

65. Zhang, M., Chen, B. (2011). *Understanding Emerging Commuting Trends in a Weekly Travel Decision Frame--Implications for Mega*

Region Transportation Planning September 2011 Ming Zhang and Binbin Chen Report 161127 Center for Transportation Research University of Texas at Austin Southw (Vol. 7).

66. Zheng, Y., Liu, L., Wang, L., & Xie, X. (2008). Learning Transportation Mode from Raw GPS Data for Geographic Applications on the Web, 247-256.